SKEPTICISM, EXTERNALISM AND
THE NATURE OF THE MIND-WORLD RELATION

BY

GAYLE BROWN

ACKNOWLEDGMENTS

TABLE OF CONTENTS

SKEPTICISM, EXTERNALISM AND
THE NATURE OF THE MIND-WORLD RELATION

By
Gayle Brown

August, 1999

Chairperson: Kirk Ludwig
Major Department: Philosophy

Externalism about thought content is the view that the identity conditions for the contents of propositional attitudes include facts about the external world and one's causal interaction with the world. This view seems to offer a promising new strategy for dealing with the problem of skepticism about the external world. The promise, however, is empty.

Argumentatively, this dissertation is divided into two main parts. The first part aims to clarify the skeptical challenge and to set down some plausible success conditions on responses to the challenge. Two points are important. First, I argue that the skeptic is committed to the claim that the mind is logically independent of the external world. Since the externalist aims to deny this claim, externalism presents us with the right sort of response to the skeptical challenge. Second, I argue that an adequate response to the skeptical

challenge must not appeal to empirical considerations. If externalism is to meet the challenge, arguments for this position must be *a priori* in nature.

In the second main part of the project, I examine the views of three prominant externalists: Hilary Putnam, Donald Davidson and Tyler Burge. I argue that the views these philosophers propose are false (or, at least, unsubstantiated) and that even if they were true, they would be of no help against the skeptic. Specifically, I argue that the most we can get from externalism is a conclusion of conditional form: if we live in a world in which such relational facts obtain, then (if externalism were true), it would follow that relational facts are relevant to the determination of thought content. Yet to suppose that we live in a world in which such relational facts obtain is to presuppose that we already know certain facts about the external world--i.e., that it is not the case that the actual world is a solipistic, spatio-temporal world. If I am correct in my critical assessment, externalist theories of thought content, as developed by their leading proponents, have no anti-skeptical force.

CHAPTER I
INTRODUCTION

Traditional philosophical reflection on the nature and extent of human knowledge has led many of us, in our more somber moments, to draw the seemingly inevitable conclusion that we do not know much of what we ordinarily take ourselves to know. This position falls under the general rubric of "philosophical skepticism." It is a position which tends to foster immediate dissatisfaction. The skeptic's negative assessment of our ordinary knowledge claims is genuinely disturbing.

It is this sense of dissatisfaction, more than anything else, I suspect, which has fueled the perennial interest in philosophical skepticism. Our intuitive conviction is that the skeptical verdict must be wrong; yet, we are forced to acknowledge that the multitude of attempts to identify and diagnose the skeptic's mistake have not met with any considerable success. Indeed, if there is any consensus to be gleaned from such endeavors, it is that they are all unsuccessful: either they fail to respect the depth of the skeptic's reasoning or they turn out, upon examination, to be less satisfying than the skeptical position itself. The search for an adequate response to the skeptical challenge continues. This dissertation takes up that search.

The goal of the project is twofold. First, I want to understand more fully the nature of the skeptical challenge; second, I want to evaluate a particular response to the skeptical challenge, one grounded in a general theory about the metaphysical or broadly logical

conditions for having thoughts about the world around us. I have in mind here attempts to argue from externalism about thought content to the falsity of skepticism about the external world. I aim to answer two questions: (1) Is externalism true?, and (2) If true, does it provide an adequate response to the skeptical challenge?

If the answer to (1) turns out to be "no" (if it turns out that externalism is false), then question (2) might appear then to be superfluous. The two questions are logically independent, of course--the second can be asked and answered independently of answering the first, and vice versa. Nonetheless, since false theories lack explanatory force, if the answer to the first question is "no," then one might very well wonder what the motivation is for pursuing the second. Since I aim to argue that externalism is false (or at least unsubstantiated), the question of why I consider the second would seem to demand some attention.

The reason is twofold. First, one would provide a deeper, more intellectually satisfying criticism of (what I shall call) the "externalist response" to the skeptical challenge by showing, not just that externalism is false, but that even if it were true (at least in its current manifestations), it would not meet the adequacy conditions on a successful response to the skeptical challenge. Moreover, it is precisely because the two questions are independent of one another that this approach would seem most promising--for if I am correct, I shall end up with two logically independent arguments against the possibility of responding to skepticism by appeal to externalism about thought content.

The second reason is somewhat more open-ended than the first. It is that by pursuing both questions we shall be able to learn something more general about what is required for

adequate response to skepticism. We shall learn to identify, and hopefully circumvent, the pitfalls that lie ahead of us. What we want is a deeper understanding of philosophical skepticism. This is why I am interested in externalism, even though I think it is false. I think that if I can pinpoint the mistake, then, perhaps, I shall be one step closer to finding a response that actually works.

How shall I establish this "deeper" criticism? How shall I establish the conclusion that even if externalism about thought content were true, it would not provide an adequate response to the skeptical challenge? I aim to establish this conclusion by showing that any attempt to resolve the skeptical challenge based on an appeal to externalist theories of thought content is unavoidably question-begging. The thrust of my criticism is that even if externalism were true, the most that could be established is a conclusion which takes the form of a conditional: if, indeed, we live in a non-solipsistic, spatio-temporal world, then relational facts are relevant to the determination of thought content. Yet to affirm the antecedent of this conditional is to beg the question at hand.

In the remainder of this introductory chapter, I provide a sketch of the argumentative strategy of the dissertation as a whole. In Chapter II, I set out the skeptical challenge in detail. Chapters III - V evaluate the anti-skeptical arguments proposed by three prominent externalists: Tyler Burge, Hilary Putnam, and Donald Davidson. Chapter VI summarizes the main conclusions of the dissertation. My conclusion is that the arguments these philosophers advance fail to provide a satisfactory response to the skeptical challenge. Let me be a bit more specific.

In Chapter II, I review a number of skeptical scenarios which have traditionally been taken to be genuine counter-possibilities to our having knowledge of the external world--e.g., the possibility that we are brains in a vat, the possibility that we are currently dreaming, and so forth. I argue that the following schema depicts the common structure of these skeptical scenarios.

(SK) [1] I can be justified in believing p only if I can rule out the possibility that q.

[2] I cannot rule out the possibility that q.

[3] Therefore, I cannot be justified in believing p.

where 'p' is replaced by a sentence about the domain in question (the external world, other minds, etc.), and 'q' is replaced by a sentence about some counter-possibility to our being justified in believing facts in that domain (the possibility that we are brains in a vat, that we are dreaming, etc.).

This observation is important for establishing my thesis in that it places me in a position to show that all arguments for skepticism rest on the same fundamental assumptions. They are the following.

(A) We can and do have *a priori* knowledge.

(C) For any subject, S, and any proposition, p, if S knows that p and S knows that p entails ~ q, then S is in an epistemic position to know that ~ q.

(I) S knows that p only if S is in an epistemic position to construct an argument on the basis of the evidence to which S has access to the conclusion that p is more likely true than false.

(E)   The only evidence S has for propositions about the external world consists of propositions about S's own mind (and *a priori* truths).

(M)   There is no (*a priori* discoverable) necessary connection between the contents of our thoughts and the actual state of our environment.

I argue that (C) is crucial to justifying appropriate instances clause [1] of SK, and that (I), (E) and (M) are crucial to justifying appropriate instances clause [2], and (A) is a presupposition of any argument which seeks to undercut our knowledge of the world around us.

Let me begin with this last assumption. I take it to be an obvious ground rule that an argument for skepticism can be taken seriously only if the premises of the argument do not fall within the scope of the conclusion. This is just to observe that the skeptic, if she is to be taken seriously, must give an argument for her conclusion. In advancing an argument--any argument--one must take oneself to be justified in believing the premises of the argument. So, the skeptic must take herself to be warranted in asserting the premises of her argument. The conclusion of the skeptic's argument is that we are not justified in believing propositions about the external world. Therefore, such propositions cannot serve as premises in the argument for skepticism or as justifiers for them. I take this to show that the skeptic is committed to the claim that we can and do have *a priori* knowledge and, in particular, that we have *a priori* knowledge of the assumptions underlying philosophical skepticism.

Let us turn to the assumption that knowledge is closed under known entailment, assumption (C). Note that as I have written it out, (C) is much weaker than the traditional formulation of the closure principle, which states that given that S knows that p and that S knows that p entails q, then S knows that q. The traditional formulation of the closure

principle has been the focal point of much apt criticism. I do not want to pause to consider such criticisms. My job here is to give what I think is the weakest formulation of the closure principle that will work for the skeptic. And this is what I have done. (C) just says that you are in a position to know all the things that you know follow from what you know, not that you have taken advantage of being in this position.

This assumption is relied on to justify instances of clause [1] of SK. According to clause [1], I can be justified in believing p only if I can rule out the possibility that q, where, let us say, 'p' is replaced by the sentence "I am presently seated in front of my computer typing away" and 'q' is replaced by the sentence "I am a brain in a vat." According to (C), if I know that I am presently seated in front of my computer and I know that my being so-situated entails that I am not a brain in a vat, then it follows that I know that I am not a brain in a vat.

Now, what clause [2] says is that I don't know that I'm not a brain in a vat. Clause [2] is justified by appeal to (I), (E) and (M). Basically, (I) says what has to happen in order for me to be in an epistemic position to know a given proposition--i.e., it sets down a necessary condition on my being in a position to know that I'm not a brain in a vat. Assumptions (E) and (M) secure the claim that I cannot meet this condition.

According to (E), the only evidence to which I have access (in ruling out the possibility that I am currently a brain in a vat) is the contents of my own thoughts (and *a priori* truths). According to (M), no proposition about the contents of my own thoughts entails any proposition about the external world. It follows from the conjunction of (E) and (M) that the evidence I have at my disposal is not sufficient to construct an argument on the

basis of evidence to which S has access to the conclusion that "I am not a brain in a vat" is false.

So, from (C) we get the claim that if I know that I am presently seated in front of my computer typing away, then I know that I am not a brain in a vat. By transposition, we may infer that if I do not know that I am a brain in a vat, then I do not know that I am presently seated in front of my computer typing away. All (C) says is that I have to rule out the possibility that a given skeptical scenario obtains. Assumption (I), an internalist account of knowledge, says something about what has to happen in order for me to rule out such possibilities; it says that I have to be able to construct an argument on the basis of the evidence I have available to me to the conclusion that the proposition expressed by the sentence "I am a brain in a vat" is (minimally) more likely to be false than true. From (E) and (M) it follows that I cannot meet (I)--that I cannot rule out the possibility that I am a brain in a vat. This secures the appropriate instances of [2]. By a simple application of modus ponens it follows that I do not know that I am presently seated in front of my computer typing away.

If my analysis of the assumptions underlying the problem of philosophical skepticism is correct, then the logical space of solutions is quite narrow. There are only five basic responses to the traditional problem of philosophical skepticism, each corresponding to the rejection of one of the above assumptions. Interestingly enough, each of the traditional responses to the skeptical challenge can be taxonomized in terms of the rejection of one of the above assumptions. I examine a sample of these traditional responses.

To reject assumption (A) is to uphold a form of naturalized epistemology, the most prominent advocate of which is Quine.[1] Few philosophers have explored the option of rejecting assumption (C). One exception here is Robert Nozick.[2] Several philosophers have sought to respond to the skeptic by denying (I). These philosophers argue that an internalist account of knowledge is mistaken and instead argue that not all the factors which go into the justification of a belief need to be "internal" to the believer. This view is known as "epistemological externalism."[3] As an example of this view, we shall take up Alvan Goldman's historical reliabilism.[4] Finally, to reject (E) is to advocate a form of Direct Realism. This line has been pursued most carefully by David Armstrong.[5]

I argue that none of these responses to the skeptical challenge are adequate. This is important for my thesis in that (a) in seeing why these responses fail, I will be in a position to set down some plausible success conditions on responses to the skeptical challenge, and (b) if I am correct in my claim that these responses do indeed fail and if the taxonomy of the logical space of solutions to the skeptical problem that I propose is indeed exhaustive, then

---

[1] Quine, W.V.O., "Epistemology Naturalized," in *Ontological Relativity and Other Essays* (New York: Columbia University Press), 1969.

[2] Robert Nozick, *Philosophical Explanations* (Cambridge: Harvard University Press, 1981), pp. 167-290.

[3] Epistemological externalism is a position which should not be confused with externalism about thought content.

[4] A.I. Goldman, "What is Justified Belief?" in Pappas (ed.), *Justification and Knowledge* (Dordrecht: Reidel, 1979), pp. 1-23.

[5] David Armstrong, *Belief, Truth and Knowledge* (Cambridge: Cambridge University Press, 1973).

it follows that the only hope for resolving the problem of philosophical skepticism rests on the plausibility of denying (M).

This option has never seemed very plausible. To reject (M) is to hold that there is a necessary connection between the contents of our propositional attitudes and the actual state of our environment. (In seventeenth and eighteenth century terminology, to reject (M) is to hold that there is a necessary connection between our "ideas" and the external world). This was Berkeley's move,[6] and it has seldom been regarded as a good one, since it fails to preserve the objectivity of the world of which we are said to have knowledge.

Like idealism, externalist theories of thought content aim to discharge the threat of skepticism by rejecting (M). But, instead of making the world depend on the mind, the externalist aims to make the mind depend on the world. Moreover, the externalist aims to do this in a manner which preserves the objectivity of the world of which we are supposed to have knowledge. On the face of it, then, externalist theories of thought content offer the most promising strategy for dealing with the problem of philosophical skepticism. However, I shall argue (in Chapters III - V) that the promise is empty.

In Chapter III, we begin our examination of the externalist response to the skeptical challenge. Chapter III takes up the views of Tyler Burge. We begin with an analysis of his landmark paper, "Individualism and the Mental." There Burge argues for a position which he labels "anti-individualism" and which has since come to be known as "social external-ism." The thrust of his position is that the identity conditions for the contents of our

---

[6] George Berkeley, *A Treatise Concerning the Principles of Human Understanding* and *Three Dialogues between Hylas and Philonous* (1713), reprinted in M.R. Ayers (ed.) *Berkeley: Philosophical Works* (London: Dent, 1975).

propositional attitudes include facts about the linguistic conventions of our community. I argue (1) that Burge's conclusion leads to a contradiction and must accordingly be rejected and (2) the most Burge could establish on the basis of social externalism is a conclusion of conditional form: if, indeed, we live in a non-solipsistic world, then linguistic conventions may be relevant to the determination of thought content.

Following this, we shall take up some of Burge's more recent work. In particular, we shall take up the anti-skeptical argument he advances in "Cartesian Error and the Objectivity of Perception." His argument proceeds in two stages: in the first stage, he argues that the traditional "Cartesian" thought experiments fail to establish what he calls the "individualist assumption"; in the second stage, he argues that without this assumption, there is no skeptical conclusion to be drawn. However, Burge's argument is seriously flawed. The individualist assumption (as he understands it) is the assumption that we *would* have the very same thoughts we have now, even if the world were radically different. However, if my analysis of the assumptions underlying philosophical skepticism is correct, then it follows that the skeptic is not committed to the individualist assumption, but only to a much weaker assumption--namely, to the assumption that we *could* have the very same thoughts we have now, even if the world were radically different.

In chapter IV, we take up the views of Hilary Putnam, beginning with his famous Twin Earth thought experiment in "The Meaning of 'Meaning.'"[7] We are interested in this paper insofar as it seems to provide an *a priori* grounding for the causal theory of reference

---

[7] Hilary Putnam, "The Meaning of 'Meaning'" *Minnesota Studies in Philosophy*, VII, K. Gunderson (ed.), Minneapolis: University of Minnesota Press, 1975; Reprinted in *Mind, Language, and Reality*, Cambridge: Cambridge University Press, 1975, pp. 215-271.

(CTR). We then turn to the anti-skeptical argument he offers in the first chapter of *Reason, Truth and History*.[8] There he argues from the truth of CTR to the falsity of the claim that we are brains in a vat (of a rather specific sort). His argument is this: given the way reference is fixed, we can have a belief (or other propositional attitude) about objects of type x (generally speaking) only if we have had the "appropriate" causal contact with tokens of x.

The "only if" has the strength of metaphysical necessity. In other words, it is a metaphysically necessary condition on our referring to and representing a given type object, that we have had causal contact with tokens of that object. However, in a later paper, Putnam relinquishes the notion of metaphysical necessity, and opts instead to talk of physical necessity.[9] Given this concession, it is apparent that the appeal to CTR in his anti-skeptical argument simply begs the question at hand. More significantly, even if Putnam's account were correct, the most he could establish is the claim that the contents of our thoughts are individuated by external factors relative to a world, only if we live in a world that admits of such factors--that is, only if we live in a physically possible world.

In chapter V, we take up the views of Donald Davidson. He begins by taking the standpoint of the radical interpreter as methodologically basic. From this standpoint, Davidson claims to have established that radical skepticism about the world is

---

[8] Hilary Putnam, "Brains in a Vat," *Reason, Truth and History* (Cambridge: Cambridge University Press, 1981), pp. 1-21.

[9] Hilary Putnam, "Is Water Necessarily H₂O?" *The Philosophy of A.J. Ayer,* L.E. Hahn, ed., (La Salle, IL: Open Court, 1992), pp. 429-454.

unintelligible.[10] He has a number of arguments for this position, the most famous of which is his Omniscient Interpreter Argument. I argue that we can take Davidson's talk of interpretation to be relevant to the problem of skepticism about the external world, only if we take him to be giving an analysis of the conditions under which interpretation is possible. To generate a valid argument against the skeptic, however, Davidson must claim that we satisfy these conditions. Yet, in arguing this claim, he would have to appeal to empirical considerations. Therefore, Davidson's discussion of interpretation is impotent against the skeptic.

Chapter VI summarizes the main conclusions of the dissertation. The most significant conclusion is that arguments for externalism tacitly appeal to facts about the external world. This violates one of the fundamental success conditions on responses to the skeptical challenge (identified in Chapter II). I conclude that to argue from externalism (in its present forms) to the falsity of skepticism about the external world is to beg the question against the skeptic.

---

[10] Donald Davidson, "The Method of Truth in Metaphysics," in *Inquiries into Truth and Interpretation* (Oxford: Oxford University Press, 1984); "A Coherence Theory of Truth and Knowledge," in *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, Ernest LePore, ed., (New York: Basil Blackwell, 1986).

CHAPTER II
SKEPTICISM ABOUT THE EXTERNAL WORLD:
THE TRADITIONAL EPISTEMOLOGICAL PROBLEM[1]

I. Introduction

The problem of philosophical skepticism has played a central role in epistemological theorizing at least since the time of Descartes, and arguably since the dawn of philosophy. Indeed, I would go so far as to say that the majority of philosophers at least tacitly agree that one measure of the adequacy of an epistemological theory is how well it handles the skeptical challenge. No doubt there is something to this line of thinking. Nonetheless, it obscures an important observation--namely, that skepticism is, itself, an epistemological theory. We might think that skepticism is a bad theory. However, that remains to be shown. It is illegitimate to rule skepticism out by *fiat*--i.e., by requiring its rejection as a criterion of adequacy on a theory of knowledge. The fact that philosophical skepticism has withstood a barrage of criticism throughout the history of philosophy testifies to the depth of the issues which prompt the skeptical verdict.

My aim in this chapter is to understand more fully the nature of the skeptical challenge. The chapter is divided into six sections. In the next section, I review a number of skeptical scenarios, taking as my starting point Descartes' *First Meditation* descent into

---

[1] My discussion of the skeptical challenge and the underlying assumptions of skepticism owes much to K. Ludwig, unpublished lecture notes.

radical skepticism. I argue that these skeptical scenarios all exhibit the same underlying structure, identified in Chapter I as (SK). Section III contains my analysis of the assumptions underlying philosophical skepticism, identified in Chapter I as (A), (C), (E), (M), and (I). I examine the role these assumptions play in generating the skeptical conclusion and set out the argument for skepticism in explicit form.

In section IV, I provide an exhaustive taxonomy of the logical space of solutions to the problem of skepticism. I argue that each of the traditional responses to the skeptical challenge can be taxonomized in terms of the rejection of one (or more) of these assumptions. I examine four representative responses in some detail: (1) Quine's rejection of (A), (2) Nozick's rejection of (C), Armstrong's rejection of (E) and Goldman's rejection of (I). I argue that these responses to the skeptical problem are inadequate.

There are two reasons why my discussion in section IV is particularly important for establishing the thesis of the dissertation. First, the success conditions I set down on responses to the challenge fall directly out of my examination of the inadequacies of these attempts to reject (A), (C), (E) and (I). I set out these success conditions in section V of this chapter. Second, if the taxonomy of solutions to the challenge that I propose is indeed exhaustive, it follows that the only remaining anti-skeptical route is by way of the rejection of (M).

As noted in Chapter I, however, this has seldom been viewed as a promising anti-skeptical strategy, since to reject (M), it has been thought, is to commit oneself to idealism-- taking this to cover any view which holds that propositions about the external world are logically dependent on propositions about minds, that is, any view that holds the external

world is mind-dependent. Traditionally, philosophers have rejected idealism because it fails to preserve the objectivity of the world about which we are said to have knowledge. The externalist response seems especially promising because it aims to reject (M), while retaining the objectivity of the world our thoughts are about. In the sixth and final section of this chapter, I briefly introduce the externalist response. This will serve as a bit of stage setting for the next three chapters of the dissertation, where I argue that the externalist response fails to meet the success conditions proposed in section V.

II. Skeptical Scenarios Reviewed

In this section, I will use Descartes' *First Meditation* arguments as a foil for introducing the problem of skepticism about the external world. I wish to emphasize at the outset that my interest is not primarily interpretative, but argumentative. I will consider the development of Descartes's skeptical reflections in the *First Meditation*. For the most part, my account will follow standard interpretations. I will note where I touch on matters of controversy, without, however, trying to resolve them decisively. It is not my intention to break new ground in the interpretive literature on Descartes. This serves merely as a stage for an evaluation of he materials that Descartes introduces in pursuit of a more radical skepticism than any for which, as I think, he aimed.

In particular, while Descartes was arguably concerned with whether we could have *certain* knowledge, as I have already indicated the sort of skepticism I am concerned with denies that we can have even minimally justified beliefs about the external world. The connection of Descartes's concerns with this concern is that although he did not intend that

his arguments have such a radical conclusion, the sorts of examples that he introduces have been thought to provide the materials for a much more powerful form of skepticism (and often enough the more powerful form has been attributed to Descartes). I am primarily concerned with that use of the considerations he advances, and the form those and allied arguments exhibit, not with whether that was Descartes's own use (as I think it was not).

Argumentatively, the primary objective of this section is to show that the following argument schema (SK) depicts the general structure of all of the skeptical arguments,

(SK) [1]   I can be justified in believing p only if I can rule out the possibility that q.

  [2]   I cannot rule out the possibility that q.

  [3]   Therefore, I cannot be justified in believing p.

where 'p' is replaced by a sentence about some domain (the external world, other minds, the future, etc.), and 'q' is replaced by a sentence about some counter-possibility to our being justified in believing facts in that domain (the possibility that we are dreaming, that we are currently trapped in "virtual reality," that there exists an evil genius who is currently deceiving us, that we are presently brains in a vat, etc.). If we can establish this claim, then we shall be in a position to show that all arguments for skepticism rely on the same fundamental assumptions.

In the opening paragraph of the *First Meditation*,[2] Descartes tells us that he has discovered that many of the beliefs he acquired in childhood have turned out to be false, and that all the beliefs which he has formed on the basis of these false beliefs are themselves doubtful. This shows, Descartes supposes, that our methods of acquiring beliefs are not reliable, since we have reason to believe that at least some of them have led to false beliefs in the past. The purpose of the skeptical arguments that Descartes considers in the *First Meditation* is to determine which of our methods of acquiring beliefs are reliable and which are not.

Descartes has two primary reasons for focusing on our methods of acquiring beliefs. The first is that he is well aware that it would be nearly impossible for him to doubt each of his beliefs, one-by-one. This, he says, "would be an endless undertaking." The second reason for adopting this strategy is that by examining our methods of acquiring beliefs, he is able to evaluate large portions of his belief system, in one fell swoop, as it were. This is clear from the following passage.

> [I]nasmuch as reason already persuades me that I ought no less carefully to withhold my assent from matters which are not entirely certain and indubitable than from those which appear to me manifestly to be false, if I am able to find in each one some reason to doubt, this will suffice to justify my rejecting the whole. And for that end it will not be requisite that I should examine each in particular, which would be an endless undertaking; for owing to the fact that the destruction of the foundations of necessity brings with it the downfall of the rest of the edifice, I shall only in the first place attack those principles upon which all my former opinions rested.[3]

---

[2] My interpretation of the First Meditation is heavily influenced by the H. Frankfurt, *Demons, Dreamers and Madmen: The Defense of Reason in Descartes' Meditations*.

[3] E.S. Haldane and G.T.R. Ross, *The Philosophical Works of Descartes*, p. 145.

We also discover, from this passage, that Descartes' method leaves intact only those beliefs which are immune from doubt. Thus, his criterion of knowledge is absolute certainty. Any method of acquiring beliefs which produces even a single belief which is open to doubt is thus denounced. Descartes is thus searching for an infallible method of acquiring true beliefs.

Descartes's first proposal for an infallible method of acquiring beliefs is this: beliefs acquired on the basis of the senses are infallible. He says, "All that up to the present time I have accepted as most true and certain I have learned either from the senses or through the senses."[4] Yet, this cannot be correct, since "it is sometimes proved to me that these senses are deceptive, and it is wiser not to trust entirely to any thing by which we have once been deceived."[5] Although in this passage Descartes does not provide any examples of cases in which the senses are deceptive, it is clear from the next passage that he has in mind "things that are hardly perceptible, or very far away."

Consequently, Descartes considers the possibility that we can be certain of those beliefs which are formed on the basis of sensory perceptions that occur under ideal conditions. He says, "But it may be that although the senses sometimes deceive us concerning things which are hardly perceptible, or very far away, there are yet many others to be met with to which we cannot reasonably have any doubt."[6] Examples of such privileged beliefs include "the fact that I am here, seated by the fire, attired in a dressing

---

[4] *Ibid.*, p. 145.

[5] *Ibid.*, p. 145.

[6] *Ibid.*, p. 145.

gown, having this paper in my hands" and so forth.[7] Descartes holds that he would be mad to doubt such beliefs.

So, Descartes supposes, for the moment, that we can be certain of those judgments which are formed on the basis of sensory perceptions which occur under ideal conditions. In the very next passage, however, Descartes introduces the Dream Argument which is intended to show that even this more refined method of acquiring beliefs is not infallible. Famously, he sets out the argument as follows:

> How often has it happened to me that in the night I dreamt that I found myself in this particular place, that I was dressed and seated near the fire, whilst in reality I was lying undressed in bed! At this moment it does indeed seem to me that it is with eyes awake that I am looking at this paper; that this head which I move is not asleep, that it is deliberately and of set purpose that I extend my hand and perceive it; what happens in sleep does not appear so clear nor so distinct as does all this. But in thinking over this I remind myself that on many occasions I have in sleep been deceived by similar illusions, and in dwelling carefully on this reflection I see so manifestly that there are no certain indications by which we may clearly distinguish wakefulness from sleep that I am lost in astonishment. And my astonishment is such that it is almost capable of persuading me that I now dream.[8]

The Dream Argument is only intended to show that even under ideal circumstances, we cannot be certain of those beliefs which are formed on the basis of sensory experience; that is, even under the best of conditions, the senses are not infallible.[9] The claim, then, is that for

---

[7] *Ibid.*, p. 145.

[8] *Ibid.*, p. 146.

[9] Let me pause for a moment to note that this interpretation is somewhat controversial. On (what might be called) the "standard interpretation" of the Dream Argument, it is thought that Descartes is calling into question the existence of the external world. John Cottingham, for example, remarks:

> Observing that there are 'no conclusive signs' by which being awake can be

any experience E, we cannot be certain that E is veridical because it is possible that E may

be produced by a dream state and not by sensory perception.

It is at this point that I want to depart from strict attention to Descartes's text and aims

to consider more generally what the philosophical significance of the sorts of considerations

he advances are. Descartes, as I have said, is interested in what we can know with certainty.

And he argues in the First Meditation on the basis of the Dream Argument that we cannot

arrive at certain knowledge merely by relying on sensory experiences. But I am interested

in the question of whether the possibility of like-like dreams shows something stronger--

namely, that we cannot have any reason whatsoever for believing anything about the external

world on the basis of our current sensory experience.

I believe that Descartes's considerations about dreaming do force us to reconsider the

claim even that we are mimimally justified in believing what we do about the external world.

From the mere fact that it seems to me that I am at this very moment awake, seated in front

---

distinguished from being asleep, Descartes proceeds, in effect, to mount a general
doubt about whether we are justified in asserting the real extra-mental existence of
any particular object which we appear to perceive via the senses. (In J. Dancy and
E. Sosa (eds.), *A Companion to Epistemology*, p. 94.)

Now, I think this is the wrong way to understand the force of the Dream Argument. For as
Cottingham himself notes, the argument, thus understood, is open to an obvious objection.

Critics of this argument have suggested that the very concept of dreaming is parasitic
on the concept of waking life, so that, again, we have not been offered a general
reason for doubting the existence of external objects. (*Ibid.*)

The difficulty here goes beyond this objection. If the Dream Argument were intended to call into
question our knowledge of (or justified belief in) the existence of the external world, then
Descartes's would not be in the position to note that he had been misled in the past; for positively
knowing that one has been misled in the past is tantamount to knowing some fact about the external
world.

of my computer, it does not follow that I am, in fact, awake and seated in front of my computer. Moreover, even if I happened to fall asleep in front of my computer, as sometimes happens, and were dreaming that I was so situated, I would not be justified in believing that I was seated in front of my computer. Dreams are not a reliable means of acquiring beliefs about the world.

What this shows has important implications--namely, that there is no necessary connection between things seeming a certain way to me and things actually being that way. I may say to myself "I know I am awake because I am now pinching my arm." But it is certainly possible that I am merely dreaming that I am now pinching my arm. This suggests that there is no way for me to rule out the possibility that I am presently dreaming on the basis of my current sensory experience, since any experience to which I may appeal in convincing myself that I am now awake seems consistent with my merely dreaming that those conditions obtain. This suggests that the problem is not just with whether I can be certain, but with whether I have any reason at all.

But why do I need to rule out the possibility that I am now dreaming? Must I know (or be justified in believing) that I am not presently dreaming in order to know (or be justified in believing) that I am seated in front of my computer right now? In other words, is the possibility that I am presently dreaming a genuine counter-possibility to my knowing (or being justified in believing) some fact about the external world? Let us consider an ordinary example.

Suppose someone walks into the office and says, "where did you park your car this morning?" I answer, "I parked over at so-and-so's apartment complex. I haven't gotten a

parking decal from the university yet." Here we are inclined to say that I know where my car is parked--it's in the parking lot at my friend's apartment complex. However, it is quite possible that my car has been towed. Now, clearly, if my car has been towed, then it is false that my car is in that lot. So, the possibility that my car has been towed is a genuine counter-possibility to my knowing that my car is parked in the parking lot at my friend's apartment complex. Unless I can rule out the possibility that my car has been towed, I cannot be said to know that my car is still in that lot.

Generally, a counter-possibility to a claim to know is some possibility which if it obtained would suffice for one's not knowing what one claims to know. On this understanding, my car's being towed is a genuine counter-possibility to my knowing it is in the parking lot, and my dreaming is a genuine counter-possibility to my knowing anything about the external world. While dreaming is not incompatible with what one dreams being so, that ne is dreaming it prevents one from thereby knowing it, since beliefs arrived at on the basis of dreams have no essential connection with their truth or even the likelihood of their being true.

It might be thought that this is not enough for a possibility to be a genuine counter-possibility.[10] For example, we would not usually pay attention to some suggested counter-possibilities to our claims to know. If someone suggested that my car had been confiscated by aliens, I would not pay the suggestion any attention or regard it as a threat to my claim to know that my car is in the parking lot where I left it. So, it might be said, not all possibilities

---

[10] See Fred Dretske, "Epistemic Operators" and "The Epistemology of Belief."

that are incompatible with what we know need to be ruled out, and so not all are genuine counter-possibilities. Perhaps the dream possibility is like this.

But this objection will not withstand examination. It is not that my car's being confiscated is not a genuine counter-possibility that explains why I don't pay attention to it, it is that I take myself to have good evidence without doing any further investigation that this possibility does not obtain. I know first of all that cars typically do stay where one parks them. It is unlikely that the cars is missing at all given that I parked it in a certain place. Secondly, I have good inductive evidence that when cars go missing, it is for a number of specific reasons which do not include alien confiscations: cars are stolen by people or they are towed by parking enforcement agencies, or the like. If I can trust what I take to be my general background knowledge of the workings of the world, then I do not have to do any further work to dismiss the proposed counter-possibility, though it is a genuine one nonetheless.

It might be thought that we could dismiss the dream possibility in the same way. Surely, we know that it is unlikely that we are dreaming when we are, for example, typing. But this is not much help, since of course to know we are typing we have to know we are not dreaming. Moreover, as we will eventually see, it is going to be difficult to limit the scope of the skeptical possibility to just present experiences.

Where, then, does this leave us? Thus far, we have made two points. The first is that the dream possibility is a genuine counter-possibility to our knowing that any particular experience is veridical. This means apparently that unless we can rule out the possibility that we are currently dreaming, we cannot be certain of any belief which is formed on the basis

sensory experience. The second is that we cannot rule out the possibility that we are currently dreaming. These are the two basic premises of the Dream Argument, which, we can note, is an instantiation of schems (SK).

[1]   I can be justified in believing facts about the external world on the basis of my current sensory experience only if I can rule out the possibility that I am currently dreaming.

[2]   I cannot rule out the possibility that I am currently dreaming.

[3]   Therefore, I cannot be justified in believing facts about the external world on the basis of my current sensory experience.

At this point, I want to introduce an argument which I take to parallel the Dream Argument. This I will call the "Argument from Virtual Reality." Consideration of this argument will show that the skeptical conclusion does not rest on particular considerations about dreaming.

Virtual reality is not merely a logical possibility; it is clearly a physical possibility. Although scientists have not perfected virtual reality, it is highly probable that one day they will.[11] Let us assume, for the sake of argument, that a group of scientists has perfected virtual reality. In virtual reality, let us suppose, things appear perfectly realistic: it seems that there are people who look and act like people, it seems as though the heat of the sun is mildly pleasant, one feels exhausted after a long hike, and so forth. Everything seems real. When

---

[11]   I use the term "virtual reality" to refer to any device that is capable of inducing experiences in us (by means of the artificial stimulation of our sensory surfaces) that are indistinguishable from the first-person perspective from experiences of a spatio-temporal world. By "in virtual reality" I meaning "in a situation in which we are the subjects of a virtual reality machine.

an individual is "hooked up" to the virtual reality system, the system's operators are able to arrange things (by running any number of appropriate programs) in such a manner that the individual's subjective experiences are phenomenally indistinguishable from the experiences that the individual would have, if she were actually engaged in the sorts of activities in which she is only virtually engaged.

Of course, the popular appeal of virtual reality does not lie in the possibility that one's ordinary everyday experiences could be "duplicated" so to speak. Most people are interested in virtual reality because it would give them the opportunity to experience the highlands of Scotland one minute, and then shift to the wonders of the Orient in the next. Few would pay money to have a virtual experience of going to the grocery store, but many of us would be willing to pay for a virtual vacation to Paris (at least those of use who neither have the time nor the financial resources to actually visit Paris). As long as we were confident that the program gave a realistic presentation of Paris, we would not feel cheated. We would feel cheated, however, if we had gone to a travel agent, planned a vacation to Paris, paid for airline tickets, and then found out, later, that the travel agent had conspired with our group of scientists.

Suppose, however, that you find yourself suspecting that you have been the victim of such fraudulent activities. You are outraged at the mere thought of having been subjected to such deception. You pick up the phone, dial the travel agent's number, the secretary answers . . . you pause. . . . How ridiculous you think to yourself! To call up the travel agent and accuse her of . . . of what? . . . of sending you and your family on a virtual vacation? You hang up the phone and head out to the local pub for a beer. You arrive at the

pub and tell your friends about the day's events. Of course, they are all amused by your paranoia. You begin to feel quite silly about the whole thing. Indeed, the more you think about the matter, the more far fetched it seems that your vacation could have been a fake. After all, on virtual vacations, airlines don't lose one's luggage! You leave the pub quite amused at your philosophical pretensions.

While driving to work the next day, you begin to reflect on your thoughts from the previous day. If you had been clear headed, you would have ruled out the possibility that your vacation was fake on the basis of the fact that you would have to have remembered being hooked up to the virtual reality system, or at least being hit over then head and loosing consciousness. Yet, you clearly remember the whole string of events without interruption. A thought, which you realize has been lingering in the back of your mind until now, forces itself to the front of your consciousness: you were already hooked up to the system. You recall that you had been wavering between going to Paris and to the Caribbean. The travel agent had set up a virtual tour of both. You were so impressed with the virtual tour of Paris that you decided to buy your airline tickets on the spot. But, you think to yourself, that was two and a half months before you left for Paris. You slam on your brakes, get out of you car, and look around, dismayed. Your astonishment is enough to convince you that it is quite possible that you are trapped in virtual reality at this very moment.

Could you, if you were in this predicament, rule out the possibility that you were currently trapped in virtual reality? Recall the crucial assumption of our thought experiment: when you are "hooked up" to the virtual reality system, the system's operators are able to arrange things (by running any number of appropriate programs) in such a manner that your

subjective experiences are qualitatively indistinguishable from the experiences that you would have, if you were not so situated. There is no way to determine whether you are trapped in virtual reality or whether you are experiencing objective reality, for in both cases your subjective experience would yield nothing by which you could discriminate a virtual-experience from a real-experience. This shows, in graphic detail, that there is no sure inference to be drawn from a person's subjective experiences to the actual state of her environment. There is a chasm between our subjective experience and the nature of the external world, a chasm which, the skeptic claims, can never be bridged.

Notice, once again, that the Argument from Virtual Reality is an instantiation of schema (SK).

[1]   I can be justified in believing facts about the external world on the basis of my current sensory experience only if I can rule out the possibility that I am currently trapped in virtual reality.

[2]   I cannot rule out the possibility that I am currently trapped in virtual reality.

[3]   Therefore, I cannot be justified in believing facts about the external world on the basis of my current sensory experience.

The Argument from Virtual Reality parallels the Dream Argument, at least on the interpretation advanced above. What this shows is that the sorts of concerns which prompt skepticism are not grounded on anything particular to reflections on dreaming. The motivation for skepticism is more general than this. It rests on considerations about how we view ourselves as individual thinkers in relation to the world around us.

Note that so far neither of these arguments is intended to call into question the existence of the external world as such. If we really were currently dreaming or trapped in virtual reality, we would still have no reason to doubt that rocks and trees, tables and chairs, or at least physical objects and the like occupy regions of space. Neither argument is sufficient to cast doubt on the existence of an independently existing reality, as is sometimes suggested, because both arguments presuppose the existence of the external world. Descartes himself is perfectly willing to concede the point that dream-images are "like painted representations" of things that really exist.[12] The scenario involving virtual reality demonstrates, even more vividly than the dreaming scenario, that the skeptic is exploiting our faith in the existence of an independently existing reality. After all, the notion of a "virtual thing" could make sense only if we were willing to suppose that there are real things. The question posed by both arguments is that of whether we can know or be justified in believing that our current sensory experience is the result of "normal" interaction with ordinary, everyday objects or whether our experience is produced by a dream state or one induced by highly sophisticated computer program. The most either argument could establish is that beliefs formed on the basis of our current sensory experience are susceptible to doubt.

Indeed, Descartes took the Dream argument to be limited in scope. Immediately following the argument, he goes on to entertain a possible objection--that even if all of one's experiences were nothing but false delusions, it would still have to be granted that "the things which are represented to us in sleep are like painted representations which can only have

---

[12] *Ibid.*, p. 146.

been formed as the counterparts of something true and real."[13] So, the objection runs, I could not have the (delusory) experiences that, by hypothesis, I do have, unless the simples, out of which these experiences are formed, are themselves represented in reality--that is, unless the simples themselves exist. (Now, of course, one difficulty in evaluating this response to the Dream Argument lies in the obscurity of the notion of a simple.)

Nonetheless, Descartes goes on to develop the objection by supposing that even if we could imagine something so completely novel--such that even the simples out of which it is composed are not represented in reality--at least "the colors of which this is composed are necessarily real."[14] From this observation, he draws the conclusion that "although these general things, to wit, [a body], eyes, a head, hands, and such like, may be imaginary, we are bound at the same time to confess that there are at least some other objects yet more simple and more universal, which are real and true; and of these just in the same way as with certain real colors, all these images of things which dwell in our thoughts, whether true and real or false and fantastic, are formed."[15]

It is in virtue of the limited scope of the Dream Argument that Descartes proposes the Evil Genius Hypothesis. He begins by contemplating the possibility that God has "brought it to pass that there is no earth, no heaven, no body, no magnitude, no place, and that nevertheless [I possess the perceptions of all these things and that] they seem to me to exist just exactly as I now see them."[16] The claim, then, is that perhaps God has destroyed (or,

---

[13] *Ibid.*, p. 146.

[14] *Ibid.*, p. 146.

[15] *Ibid.*, p. 146.

[16] *Ibid.*, p. 146.

perhaps, had never created) all the furniture of the Earth, and that what he (Descartes) is currently perceiving is nothing but a false delusion. However, Descartes is understandably reluctant to speak of God as an arch deceiver. For this reason, he introduces the idea of a supremely powerful evil genius who is bent on deceiving him at every turn.

> I shall then suppose, not that God who is supremely good and the fountain of truth, but some evil genius not less powerful than deceitful, has employed his whole energies in deceiving me; I shall consider that the heavens, the earth, colors, figures, sound, and all other external things are nought but the illusions and dreams of which this genius has availed himself in order to lay traps for my credulity; I shall consider myself as having no hands, no eyes, no flesh, no blood, nor any senses, yet falsely believing myself to possess all these things . . . .[17]

In this passage, Descartes is entertaining the possibility that all of his thoughts have been placed in his mind by of the arch deceiver. This hypothesis leads him to suppose that everything presented to his mind is nothing but "illusions and dreams."

Now, in our consideration of the dream argument we saw that Descartes was willing to concede the point that the images produced by dreams may be like painted representations of things that really exist. At that point in the Meditations, it looks as though Descartes is forced to concede this point, since he has no other explanation for how the mind came to be furnished with such images. (Of course, the move here is purely dialectic; the concession sets the stage for a discussion of the Evil Genius Hypothesis.) The possibility that there is an evil genius placing ideas in his mind is Descartes' alternative proposed explanation.

---

[17] *Ibid.*, p. 148.

The Evil Genius hypothesis is intended to show that the thoughts we would have if there were an evil genius could be qualitatively indistinguishable from the thoughts we would have if there were a common sense world. From such an impoverished epistemic perspective, we cannot rule out the possibility that there is an evil genius bent on deceiving us. Again we can note that the Evil Genius Hypothesis is an instantiation of schema (SK).

[1]   I can be justified in believing facts about the external world on the basis of my current sensory experience only if I can rule out the possibility that there is an evil genius.

[2]   I cannot rule out the possibility that there is an evil genius.

[3]   Therefore, I cannot be justified in believing facts about the external world on the basis of my current sensory experience.

Before pausing to consider the assumptions implicit in this argument, it will be useful to examine another skeptical scenario, one which is often heralded as a contemporary version of the Evil Genius Hypothesis--the brain in a vat scenario.

In this scenario, we are asked to imagine an individual who (unbeknownst to him) has been subjected to an operation in which an evil scientist has removed his brain from his body and placed it in a vat of nutrients. We then suppose that the brain's nerve endings are connected to a highly sophisticated computer system. By running an appropriate program, the scientist is able to obliterate all of the subject's memories of the brain operation and induce the experiences necessary for a continuation of consciousness. The result is that from a purely phenomenal standpoint, the subject's life seems to proceed without disruption. Nonetheless, his experiences would be illusory. By running any number of programs, the

scientist could systematically manipulate the subject's experiences. Could we, if we were

in this predicament, determine that we were not brain in a vat? When we imagine ourselves

in this predicament and entertain this question, we realize that there is no way for us to

determine that we are not brains in a vat. The Brain in a Vat Argument is an instantiation

of schema (SK).

[1]   I can be justified in believing facts about the external world on the basis of
       my current sensory experience only if I can rule out the possibility that I am
       a brain in a vat.

[2]   I cannot rule out the possibility that I am a brain in a vat.

[3]   Therefore, I cannot be justified in believing facts about the external world on
       the basis of my current sensory experience.

Thus far we have considered a number of skeptical scenarios which seem to undermine our

ordinary knowledge claims about the external world. We have found that they all exhibit

a similar structure:

(SK)   [1]   I can be justified in believing p on the basis of my current sensory experience
              only if I can rule out the possibility that q.

       [2]   I can't rule out the possibility that q.

       [3]   Therefore, I cannot be justified in believing p on the basis of my current
              sensory experience.

This is an important conclusion. The fact that the various skeptical scenarios exhibit a

common structure shows that the skeptical conclusion is ultimately independent of particular

considerations about, say, dreaming--that, in fact, skepticism rests on general considerations

about how we conceptualize the relation that obtains between ourselves as individual thinkers and the external world, and how we conceptualize our epistemic perspective. In the next section of this chapter, I aim to isolate the fundamental assumptions on which philosophical skepticism rests and to lay out the argument for skepticism in explicit form.

III. The Assumptions of Philosophical Skepticism

In this section, I aim to show that arguments for skepticism rest on a total of five assumptions. They are the following.

(C)   If I know that p entails ~q, then I can know p only if I am in an epistemic position to know ~q.

(I)   I can know facts in about the external world only if I am in an epistemic position to construct an argument on the basis of my evidence to the conclusion that such facts are more likely true than not.

(E)   The only evidence that I have at my disposal is the contents of my own thoughts.

(M)   There is no necessary connection between propositions about the contents of my own thoughts and propositions about the external world.

(A)   We can and do have *a priori* knowledge.

Assumption (C) is crucial to justifying clause [1] of (SK). Assumptions (I), (E) and (M) are crucial to justifying clause [2] of (SK). Assumption (A) is a presupposition of any skeptical argument.

To further clarify the role these assumptions play in generating the skeptical conclusion, let us have a look at skepticism with respect to some other domains. I want to

show that (SK) depicts the general structure of skeptical arguments which issue a challenge to our knowledge of other minds and inductive inferences. If we can show this, then we will have taken a further step in showing that all arguments for skepticism rest on the same fundamental assumptions. Let us begin by considering the problem of skepticism about other minds.

Arguments for the conclusion that we can have no knowledge or justified belief about other minds typically begin by calling attention to the apparent asymmetry between how we come to know our own mental states and how we come to know the mental states of others. Thus, the argument begins with what seems to be a brute fact about our own minds: we know what we think and feel directly and immediately. Some philosophers have even supposed that we have incorrigible access the contents of our own thoughts. When it comes to other people, however, we must *infer* what's going on in their minds on the basis of their behavior--the assumption at play here is that we do not have direct knowledge of the mental states of others.

However, there is no necessary connection between a person's exhibiting a certain behavior and that person's being in such-and-such a mental state. This is only to observe that we do make mistakes: on occasion, we observe a friend's behavior and assume that she is angry, only later to discover that she was, in fact, just tired. Now, is it not possible that our friend possesses no mental states at all, but is merely a cleverly disguised automaton? A moment's reflection suggests that, as improbable as it may be, we cannot be assured that there are any other minds in the universe; and that, for all we know, they could all be cleverly disguised automatons.

Notice, once again, that the structure of the argument is an instantiation of the argument schema (SK).

[1]   I can be justified in believing facts about other minds on the basis of the behavior of bodies only if I can rule out the possibility that they are cleverly disguised automatons.

[2]   I cannot rule out the possibility that they are cleverly disguised automatons.

[3]   Therefore, I cannot be justified in believing facts about other minds on the basis of the behavior of bodies.

Why can't we rule out the possibility that those around us are cleverly disguised automatons? We can't rule out this possibility because (E) our evidence for facts about other minds is limited to facts about their behavior and (M) no fact about their behavior entails any fact about their minds. Similarly, we can't rule out the possibility that we are currently trapped in virtual reality because (E) our evidence for facts about the external world is limited to facts about our own minds (our own subjective experiences) and (M) no fact about our own minds entails any fact about the external world.

Let us move onto the problem of induction, which once again, we can note, is an instantiation of (SK). Here we take up Hume's traditional statement of the problem. Hume begins with some general observations about science and mathematics. No future observation will refute the claim that a triangle has three sides, but this is not so for the claims of science. Science gives us general hypotheses which help us predict future events. Such hypotheses can be highly corroborated, but they can never be proven true. We can show that a hypothesis is false by observing a single counter-instance, but even if every

experience we have points toward the truth of a given hypothesis, we can never be absolutely certain that the hypothesis is true--consider, for example, the fact that Newton's so-called "Laws of Nature" have been effectively falsified.

Hume's argument begins with a distinction between what he calls "ideas" and "impressions." The primary difference between the two is a matter of "force" and "livelyness."

> By the term 'impression' then I mean all our more lively perceptions, when we hear, or see, or feel, or love, or hate, desire, or will. And impressions are distinguished from ideas, which are the less lively perceptions of which we are conscious when we reflect on any of those sensations or movements above mentioned.[18]

At first glance, it seems that the human mind is unbound and limitless. "Nothing, at first view, may seem more abounded that the thought of man, which not only escapes all human power and authority, but is not ever restrained within the limits of nature and reality."[19] Yet, Hume is going to argue that this claim is incorrect--that there are limits on what the human mind can do.

The human mind can only manipulate that which has been presented to it. For example, a person, blind from birth, cannot truly be said to entertain notions of colors. Likewise, a person born deaf cannot enjoy, say, Beethoven's Fifth Symphony. This shows, Hume claims, that where there is no impression, there can be no idea. This claim extends not only to those born blind or deaf, but the rest of us as well. Hume gives the following

---

[18] David Hume, *Enquiries Concerning Human Understanding*, §II.

[19] *Ibid.*, §II.

example: "When we think of a golden mountain, we only join two consistent ideas, 'gold' and 'mountain' with which we were formerly acquainted."[20] "We shall always find," Hume says, "that every idea which we examine [every idea that comes before the mind] is copied from a similar impression."[21]

Hume next draws a distinction between two kinds of reasoning: relations of ideas and matters of fact. Here is what Hume has to say about these two kinds of reasoning.

> Of the first kind are the sciences of Geometry, Algebra, and Arithmetic, and in short, every affirmation which is either intuitively or demonstratively certain. *That the square of the hypotenuse is equal to the square of the two sides* is a proposition which expresses a relation between figures. *That three times five is equal to the half of thirty* expresses a relation between these numbers. Propositions of this kind are discoverable by the mere operation of thought, without dependence on what is anywhere existent in the universe.[22]

In contrast, Hume says:

> Matters of fact . . . are not ascertained in the same manner, nor is our evidence of their truth, however great, of a like nature with the foregoing [i.e., relations of ideas]. The contrary of every matter of fact is still possible, because it can never imply a contradiction and is conceived by the mind with the same facility and distinctness as if ever so conformable to reality. *That the sun will not rise tomorrow* is no less intelligible than that *it will rise*. We should in vain, therefore, attempt to demonstrate its falsehood. Were it demonstrably false, it would imply a contradiction and could never be distinctly conceived by the mind.[23]

---

[20] *Ibid.*, §II.

[21] *Ibid.*, §II.

[22] *Ibid.*, §IV, I.

[23] *Ibid.*, §IV, I.

Hume's claim here is this: take any proposition whatsoever and ask yourself "can you conceive of that proposition being false." Now, if you answer "no" then you've got a proposition which falls within the class of relations of ideas. For example, consider the proposition *all circles are round*. Now, try to conceive of a circle that is not round, say, a square circle. This involves a contradiction, Hume claims. Therefore, a round square "could never be distinctly conceived in the mind."[24]

On the other hand, try to imagine *that you are not reading this dissertation right now.* Is this possible? Sure--you might have fallen asleep (after reading the first chapter) and so you might just be dreaming that you are currently reading this dissertation. More to the point, consider the proposition *that the force of attraction between two bodies is proportional to each of their masses and inversely proportional to the square of the distance between them.* (This is Newton's Law of Universal Gravitation.) Now, is it possible to conceive Newton's Law being false? Sure--it is possible that every time I leave my office, all the books on my bookshelf float in the air. This shows that Newton's Law of Universal Gravitation falls within the class of propositions about matters of fact.

Hume's important claim is this: "All reasoning concerning matters of fact seem to be founded on the relation of *cause* and *effect*."[25] Consider the case of Adam who has just appeared upon the face of the Earth. Suppose that he has no experience of the world. Could Adam know, just by looking at, say, water, that it would quench his thirst? Would Adam be able to know a priori that he would drown in the water? Could he know that water is capable

---

[24] *Ibid.*, §IV, I.

[25] *Ibid.*, §IV, I.

of extinguishing forest fires, leaving spots on cars, or keeping kittens off the kitchen counters? The answer is "no"--he could not know all these things because he has no experiences of them.

We cannot know *a priori* (independently of experience) that when I release this book from my hand that it will fall to the ground (since it is clear that we can conceive of the book just hovering in the air)--i.e., the proposition that the book will hover in the air does not involve a contradiction. But, why is it that we all think that when I release this book that it will fall to the ground? Hume's answer is that this is what we have experienced in the past-- every time someone drops an object, that object falls to the ground (unless otherwise held in place).

> As to past experience, it can be allowed to give direct and certain information of those precise objects only, and that precise period of time which fell under its cognizance: But why this experience should be extended to future times and to other objects which, for aught we know, may be only in appearance be similar, this is the main question on which I would insist.[26]

Hume's point here is this: we can only judge of things that have been presented to the mind; future and unobserved events have not been presented to the mind; therefore, we cannot judge of them. Why not? Well, to do so we would have to know the truth of some principle which bridges past experiences and future experiences: that the future will resemble the past. Yet, how do we know that the future will resemble the past? Hume's claim is that the proposition *that the future will resemble the past* can only be known on the basis of past

---

[26] *Ibid.*, §IV, I.

experience. (It is, after all, conceivable that the universe change drastically tomorrow.) Our attempt to ground matters of fact is thus circular. It follows that reasoning based on matters of fact does not yield certainty. Moreover, since circular reasoning provides no support whatsoever for its conclusion, it follows that we cannot have any reason at all to believe in conclusions drawn from matters of fact reasoning.

Once again, we can note that the problem of skepticism about induction is an instantiation of (SK).

[1]   I can be justified in believing facts about the future or unobserved events only if I can rule out the possibility that the universe will change drastically tomorrow.

[2]   I cannot rule out the possibility that the universe will change drastically tomorrow.

[3]   Therefore, I cannot be justified in believing facts about the future or the unobserved.

Hume's argument rests on the same fundamental assumptions discussed above. We cannot rule out the possibility that the universe will change drastically tomorrow because (E) our evidence for facts about the future (the unobserved) is limited to facts about the past (the observed), and (M) no fact about the past (the observed) entails any fact about the future (the unobserved).

In addition, what Hume's argument brings to the forefront, perhaps even more so than the other skeptical arguments we have considered, is the assumption that to solve the problem of skepticism, one must be in a position to give an *a priori* argument on the basis of the evidence one has at one's disposal for the conclusion that we have more reason than

not for thinking that we have knowledge in that domain. In the case of skepticism about induction, we have the two previously identified assumptions (E) that our evidence is limited to observed facts about the present and the past and (M) that no fact about the present or the past entails any fact about the future. In addition to these, we can see that the skeptic also needs to assume that we can have knowledge about the future only if we are in an epistemic position to construct an *a priori* argument which appeals only to facts about the past for the conclusion that we have more reason than not for thinking that propositions about the future are true. This is assumption (I).

Looking back on the arguments for skepticism about other minds and the external world, we can see that these same assumptions are at work. Let us return to our instantiation of (SK) with respect to the problem of other minds and examine the role these assumptions play in generating the skeptical conclusion.

(SK) [1] I can be justified in believing facts about other minds on the basis of the behavior of bodies only if I can rule out the possibility that they are cleverly disguised automatons.

[2] I cannot rule out the possibility that they are cleverly disguised automatons.

[3] Therefore, I cannot be justified in believing facts about other minds on the basis of the behavior of bodies.

Clause [1] of (SK) is justified by a sufficiently weakened formulation of the Closure Principle: if S knows that p, and S knows that p entails ~q, then S is in an epistemic position to know that ~q. (This is weaker than the traditional formulation of the Closure Principle in that it only requires that S be in an epistemic position to know that ~q and not that S actually

know that ~q.) The skeptic reasons in reverse--since S cannot know that ~q, it follows that S does not know that p.

Clause [2] follows from (I), (E), and (M). What has to happen in order for S to be able to rule out the possibility that those around her are merely cleverly disguised automatons? The skeptic's answer is that S can rule out this possibility only if she is in an epistemic position to construct an argument which appeals only to the evidence she has at hand to the conclusion that propositions about other minds are more likely true than not. This is assumption (I). Assumption (E) says that S's evidence is limited to facts about the behavior of bodies. Assumption (M) says that there is not necessary connection between propositions about the behavior of other bodies and propositions about other minds. In other words, it follows from (E) and (M) that S cannot meet (I), and it follows from (C) and [1] that unless S can meet (I), then S cannot know facts about other minds. Thus, S cannot know facts about other minds. It is significant that the argument for skepticism about other minds has the exact same structure as the argument for skepticism about the external world.

We are now in a position to state, in general terms, the master argument for philosophical skepticism. To this end, let 'p' be a sentence about the external world and let 'q' be a sentence which describes some counter-possibility to our having knowledge in that domain.

[1]   I can know that p is true only if I know that q is false. (C)

[2]   I can know that q is false only if I am in an epistemic position to construct an argument on the basis of the evidence that I have available to me for the conclusion that q is false. (I)

[3]   The only evidence to which I have access in constructing such an argument is the contents of my own thoughts. (E)

[4]   There is no necessary connection between propositions about the contents of my own thoughts and propositions about the external world. (M)

[5]   Therefore, I am not in an epistemic position to construct an argument on the basis of my evidence to the conclusion that q is false. (from 3 and 4)

[6]   Therefore, I cannot know that q is false. (from 2 and 5)

[7]   Therefore, I cannot know that p is true. (from 1 and 6)

We have arrived at a valid argument whose premises are seemingly true. Yet, the conclusion of the argument strikes us as obviously false. The problem of philosophical skepticism is thus the problem of finding reasonable grounds for rejecting at least one of the premises of the skeptic's argument. Here, there are only five options; each of these options corresponds to the rejection of one of the previously identified assumptions.

IV. The Logical Space of Solutions

Each of these assumptions has been denied at one point or another as a way of dealing with the skeptical challenge. For example, Robert Nozick in his book *Philosophical Explanations* aims to discharge the threat of skepticism by rejecting assumption (C). Quine and his fellow naturalists have issued a challenge to (A). Alvin Goldman gives a sustained attack on (I) in his paper "What is Justified Belief?" and other writings. David Armstrong's Direct Realism aims to dispel the skeptical challenge by rejecting (E). I argue that these sorts of responses are unsuccessful. If this taxonomy is indeed exhaustive, the only remaining anti-skeptical route is by way of the rejection of assumption (M). Here there are only two

options: Idealism and Externalism about Thought Content. I argue that Idealism is not a promising strategy. If my argument here is convincing, it follows that the only live option is externalism about thought content.

<u>Rejecting A: Quine's Response</u>

Contemporary discussions of *a priori* knowledge have been shaped largely by Kant, whose treatment of *a priori* knowledge centers around three main distinctions. It will be useful, at the outset of our project, to clarify these distinctions. The first is the epistemological distinction between *a priori* and *a posteriori* knowledge. The second is the metaphysical distinction between necessity and contingency. And, the third is a semantical distinction between analytic truths and synthetic truths.

To begin with the third distinction, Kant uses the terms 'analytic' and 'synthetic' to distinguish two kinds of judgments. An analytic judgment is one whose predicate adds nothing to the concept of the subject. For example, the judgment 'All bodies are extended' is analytic, since the concept of extension is contained in the concept of body. In contrast, a judgment is synthetic if in making the judgment, something is added to the concept of the subject which was not already 'thought' in the predicate. For example, the judgment 'All bodies are heavy' is synthetic, since the predicate adds something which is not already contained in the concept of body. Analytic judgments, Kant thought, are such that to deny them would involve a contradiction. This is not true of synthetic judgments.

The second distinction is that between necessity and contingency. A proposition is necessarily true just in case it couldn't possibly be false. Or, in terms of possible worlds, a

proposition is necessary just in case it is true in all possible worlds. A proposition is contingent just in case it could be falsified. Or, in other words, just in case it is false in at least one possible world.

The final distinction is that between *a priori* and *a posteriori* knowledge. The standard way of setting out this distinction is by reference to Kant's claim that *a priori* knowledge is "absolutely independent of all experience."[27] The two marks of the *a priori*, according to Kant, are necessity and universality.

> What we here require is a criterion by which to distinguish with certainty between pure and empirical knowledge. Experience teaches us that a thing is so and so, but not that it cannot be otherwise. First, then, if we have a proposition which in being thought is thought as necessary, it is an a priori judgment. . . . Secondly, experience never confers on its judgments true or strict, but only assumed and comparative universality, through induction. We can properly only say, therefore, that, so far as we have hitherto observed, there is no exception to this or that rule. If, then, a judgment is thought with strict universality, that is, in such manner that no exception is allowed as possible, it is not derived from experience, but is valid absolutely a priori.[28]

Kant's claim here is that judgment is *a priori* if it is necessary and universal, and that if a judgment is necessary and universal then it is *a priori*. Now, since Kant took it to be evident that we do know certain truths which are necessary and universal (for example, the truths of mathematics), it follows that we can and do have *a priori* knowledge.

Kant agreed with the received opinion that all analytic judgments are *a priori*. However, he denied that only analytic judgments are *a priori*, arguing that there are some

---

[27] Immanuel Kant, *Critique of Pure Reason*, p. 43.

[28] *Ibid.*, pp. 43-44.

judgments which are synthetic that can be known *a priori*. The truths of mathematics and geometry, Kant thought, were synthetic *a priori* judgments.

Thus, there are three claims that Kant sought to defend. The first is that a judgment is *a priori* if and only if it is both necessary and universal. The second is that we can and do have *a priori* knowledge. And the third is we can and do have synthetic *a priori* knowledge. Kant's main argument in defense of these claims is that we could not account for the mathematical knowledge that we possess unless they were true.

The notion of the synthetic *a priori*, however, has met considerable opposition, particularly by philosophers of an empiricist bent. The problem, for the empiricist, is to account for the possibility of necessary truths. Ayer puts the problem as follows:

> [A]s Hume conclusively showed, no general proposition whose validity is subject to the test of actual experience can ever be logically certain. No matter how often it is verified in practice, there still remains the possibility that it will be confuted on some future occasion. The fact that a law has been substantiated in n - 1 cases affords no logical guarantee that it will be substantiated in the nth case also, no matter how large we take n to be. And this means that no general proposition referring to a matter of fact can ever be shown to be necessary and universally true.[29]

As Ayer, himself, notes, "if empiricism is correct no proposition which has a factual content can be necessary."[30] So, either the empiricist is forced to deny that the truths of mathematics and logic are necessary, or the empiricist must show that such necessary truths are devoid of

---

[29] A.J. Ayer, "The *A Priori*," p. 72.

[30] *Ibid.*, p. 72.

any factual content. Ayer, himself, opted for the second line, claiming that all necessary truths, including the truths of mathematics and logic, are analytic.

Ayer sought to ground *a priori* knowledge on the notion of analyticity. "The principles of logical and mathematics," Ayer wrote, "are true universally simply because we never allow them to be anything else." This is because "we cannot abandon them without contradicting ourselves, without sinning against the rules which govern the use of language, and so making our utterances self-stultifying." As Ayer makes the point, "the truths of logical and mathematics are analytic propositions or tautologies."[31]

According to Ayer, Kant's treatment of the analytic/synthetic distinction is obscured by the use of the term 'concept'. Moreover, Kant implicitly assumes that all analytic propositions are of the subject/predicate form, which is clearly a mistake. Even waiving these objections, however, there is still a crucial error in Kant's treatment of the issue. Ayer notes that Kant actually gives two nonequivalent criteria for distinguishing analytic from synthetic propositions—a psychological criterion and a logical criterion. It is on the psychological criterion that Kant declares that mathematical propositions are synthetic. According to Ayer, Kant's "ground for holding that the proposition '7 + 5 = 12' is synthetic is . . . that the subjective intension of '7 + 5' does not comprise the subjective intension of '12'."[32] On the other hand, it is on the logical criterion that Kant claims that "'all bodies are

---

[31] *Ibid.*, p. 77.

[32] *Ibid.*, p. 78.

extended' is an analytic proposition" since "it rests on the principle of contradiction alone."[33]

The difficulty, Ayer says, is that:

> [A] proposition which is synthetic according to the former criterion [the psychological criterion] may very well be analytic according to the latter [the logical criterion]. For . . . it is possible for symbols to be synonymous without having the same intensional meaning for anyone: and accordingly from the fact that one can think of the sum of seven and five without necessarily thinking of twelve, it by no means follows that the proposition '7 + 5 = 12' can be denied without self-contradiction.[34]

Ayer proposes the following criterion which, he thinks, preserves the logical import of Kant's distinction between analytic and synthetic proposition. On this criterion, "a proposition is analytic when its validity depends solely on the definitions of the symbols it contains, and synthetic when its validity is determined by the facts of experience."[35] For example, the proposition 'There are ants which have established a system of slavery' is a synthetic proposition, since its truth of falsity cannot be determined solely in virtue of the definitions of the symbols which constitute it. In contrast, the proposition 'Either some ants are parasitic or none are' is analytic, since it can be determined to be true solely in virtue of the logical particles. Ayer's contention is that propositions of the latter sort convey no information about the world. They are devoid of factual content. And, it is for this reason that no experience could possibly confute them.

---

[33] *Ibid.*, p. 78.

[34] *Ibid.*, p. 78.

[35] *Ibid.*, p. 78.

The traditional analysis of *a priori* knowledge in terms of analyticity has been attacked, famously, by Quine.[36] The basis of Quine's complaint is that the notion of analyticity is in need of explication. On Quine's view, for any analysis of analyticity to be adequate, it must meet the following restriction:

[R] An adequate analysis of analyticity must not rely on any unexplicated notions.

Quine argues that the various formulations of the traditional account of analyticity--for example, analyses in terms of synonymy, interchangeability, semantical rules, meanings--all violate this restriction.

Let us consider Quine's case against the analysis of analyticity in terms of meanings. On this analysis, a statement is analytic just in case it is true in virtue of the meanings of the terms it contains and not in virtue of any fact about the world. This account, Quine notes, would be unproblematic if we had some way to understand the notion of meaning.

One point to note, however, is that meaning is not to be confused with naming. The expressions '9' and 'the number of planets' pick out one and the same abstract entity, but they do not have the same meaning, since astronomical observation was required (and not mere reflection on meanings) to determine the sameness of the entity in question. Quine's complaint is that appealing to the unexplicated notion of meaning in order to explicate the notion of analyticity is no analysis at all.

---

[36] W.V.O. Quine, "Two Dogmas of Empiricism," pp. 20-46.

To clarify the target of his attack, Quine observes that statements which are commonly regarded as analytic fall into two classes. In the first class, we find statements which are logically true, such as the following:

(1)   No unmarried man is married.

The distinctive mark of a logically true statement is that, holding the logical constants fixed, it remains true under all reinterpretations. It is the second class of statements with which Quine is particularly concerned; statements such as the following:

(2)   No bachelor is married.

It is commonly claimed that such statements can be transformed into logical truths by simply replacing synonyms for synonyms. So, for example, (2) can be transformed into (1) by simply replacing 'bachelor' with its synonym 'unmarried'. The analysis of analyticity in terms of synonymy, however, is insufficient, since we have been given no independent analysis of the notion of synonymy.

Let us see how Quine's argument works against Ayer's criterion for drawing the analytic/synthetic distinction. On Ayer's formulation:

[C]   A proposition is analytic when its validity depends solely on the definitions of the symbols it contains, and synthetic when its validity is determined by the facts of experience.

On Ayer's criterion, (2) is a logical truth, by definition--'bachelor' is simply defined as 'unmarried man'. Quine's objection runs as follows:

> But how do we find that 'bachelor,' for example, is defined as 'unmarried man'? Who defined it thus, and when? Are we to appeal to the nearest dictionary, and accept the lexicographer's formulation as law? Clearly this would be to put the cart before the horse. The lexicographer is an empirical scientist, whose business is the recording of antecedent facts; and if he glosses 'bachelor' as 'unmarried man' it is because of his belief that there is a relation of synonymy between those forms, implicit in general or preferred usage prior to his own work. The notion of synonymy presupposed here has still to be clarified. . . . Certainly the 'definition' which is the lexicographer's report of an observed synonymy cannot be taken as the ground of the synonymy.[37]

Thus, Quine's complaint is that Ayer's criterion violates condition R, since it implicitly relies on the unexplicated notion of synonymy. More generally, Quine's complaint is that the various analyses of *a prioricity* are inadequate, since they all appeal at some point to an unexplicated notion.

Now, notice that if we had no understanding of the notion of synonymy whatsoever, Quine's criticism would be right on target. His objection, however, is that "just what it means to affirm synonymy, just what the interconnections may be which are necessary and sufficient in order that two linguistic forms be properly describable as synonymous, is far from clear."[38]

In order for Quine's objection to come off, I think he actually needs a much stronger claim. Namely, Quine needs to show that the notion of synonymy is, in principle, unintelligible. But, in discussing stipulative definitions, Quine assumes that we have some pre-theoretic notion of synonymy. In the case of stipulative definitions, Quine claims,

---

[37] *Ibid.*, p. 24.

[38] *Ibid.*, p. 22.

the definiendum becomes synonymous with the definiens simply because it has been created expressly for the purpose of being synonymous with the definiens. Here we have a really transparent case of synonymy created by definition; would that all species of synonymy were as intelligible.[39]

At best, Quine can claim that an adequate analysis of analyticity has yet to be given, for he has not shown that the notion of synonymy is, in principle, unintelligible. This, I think, is sufficient to ward off Quine's attack on the analytic/synthetic distinction, and hence, to secure the possibility of *a priori* knowledge.

Before further evaluating this argument, I want to dismiss a charge that is often leveled at Quine's naturalized epistemology. There seems to be a growing consensus that Quine has not given us an answer to the age-old question of whether we can have knowledge or justified belief; rather, he has simply changed the question. The line of thought behind this criticism is that a satisfactory response to the skeptical challenge must not appeal to empirical considerations, for such considerations would fall within the scope of the skeptic's negative assessment of our knowledge, and hence to assert that we have such knowledge is simply to beg the question against the skeptic. Indeed, Quine seems to invite this charge. At one point, he acknowledges that

a surrender of the epistemological burden to psychology is a move that was disallowed in earlier times as circular reasoning. If the epistemologists goal is validation of the grounds of empirical science, he defeats his purpose by using psychology or other empirical science in the validation.[40]

---

[39] *Ibid.*, p. 26.

[40] *Ibid.*, pp. 75-76.

Remarking on this passage, Stroud charges that if the *surrender* of the epistemological burden to psychology (and to the other sciences) "does not defeat Quine's purpose, although it would defeat the traditional epistemologist's purpose, is that because the question Quine asks and answers is different from the traditional question?"[41] Stroud's answer to this question is "yes." I think this objection is misguided. Quine has not changed the subject. He has given us a principled reason for thinking that a general methodological assumption of philosophical skepticism--that we can and do have *a priori* knowledge--is unsubstantiated.

It is Quine's rejection of the analytic/synthetic distinction and his ensuing attack on the notion of *a prioricity* that undergirds his response to the skeptic, for with the rejection of these notions the distinction between philosophy and natural science is removed. The result is, in Quine's terms, a "naturalized epistemology." Epistemology, he says:

> is best looked upon . . . as an enterprise within natural science. Cartesian doubt is not the way to begin. Retaining our present beliefs about nature, we can still ask how we can have arrived at them. Science tells us that our only source of information about the external world is through the impact of light rays and molecules upon our sensory surfaces. Stimulated in these ways, we somehow evolve an elaborate and useful science. How do we do this, and why does the resulting science work so well? These are genuine questions, and no feigning of doubt is needed to appreciate them. They are scientific questions about a species of primates, and they are open to investigation in natural science, the very science whose acquisition is being investigated.[42]

---

[41] Barry Stroud, *The Significance of Philosophical Skepticism*, p. 223.

[42] W.V.O. Quine, "The Nature of Natural Knowledge," p. 68.

Quine is targeting the very methodology Descartes (or skeptics following in his First Meditation footsteps) employed to arrive at his negative conclusion about the whole of our knowledge.

Nonetheless, I think there is a reason to think that Quine's naturalized epistemology is an inadequate response to the traditional problem of skepticism. There is something that Quine has failed to explain. He has failed to explain or account for the intuitive appeal of the Cartesian thought experiments. If, indeed, there is no cosmic exile, as Quine claims, then we ought to be able to detect some fairly obvious mistake in the sorts of thought experiments on which skepticism rests. But neither Quine, nor his followers, have provided a diagnosis of the appeal of such thought experiments. In the next section of this chapter, I will develop this idea more fully. I want to set down as a success condition on an adequate response to the skeptical challenge that it provide a diagnosis of the skeptic's mistake, if, indeed, the skeptic has made a mistake. In other words, an adequate response ought to explain away the intuitive appeal of the Cartesian thought experiments.

<u>Rejecting C: Nozick's Response</u>

Although philosophers tend to disagree on the particular formulation of the closure principle involving known entailment, most would accept something like formulation (C). One exception here is Robert Nozick.[43] Nozick agrees with the skeptic that we could never rule out the possibility that any given skeptical scenario obtains; however, he claims that we

---

[43] Robert Nozick, *Philosophical Explanations*.

can know facts about the external world even though we cannot rule out such possibilities. This follows from his analysis of propositional knowledge.

A knows that P iff:

    (1)   P is true.

    (2)   A believes that P.

    (3)   If P were not true, A would not believe that P.

    (4)   If P were true, A would believe that P.[44]

So, for example, I know that I am sitting in front of my computer typing at the keyboard just in case I am in fact seated in front of my computer typing away, I believe that I am so situated, if I were not so situated, then I would not believe that I was, and if I were so situated, then I would believe that I am so situated. On Nozick's analysis of knowledge, I need not rule out the possibility that a given skeptical scenario obtains in order to know that I am sitting in front of my computer.

Let "SK" stand in for a proposition to the effect that this or that skeptical scenario obtains. Now, consider the application of Nozick's analysis of knowledge to the question of whether an ordinary person knows that it is not the case that SK (~SK). We get the following:

    (1)   ~SK is true

    (2)   A believes that ~SK

---

[44] *Ibid.*, pp. 172-178.

(3)   If ~SK were not true, then A would not believe that ~SK

(4)   If ~SK were not true, then A would believe ~SK.[45]

Take the possibility that we are trapped in virtual reality as an example. If we were trapped in virtual reality, we would believe that ~SK, so (3) is false. If we were trapped in virtual reality, we would believe that we were not. All this says, Nozick maintains, is that we cannot rule out such skeptical possibilities. To reach the skeptical conclusion--the conclusion that we do not know facts about the external world-- the skeptic must assume that knowledge is closed under known entailment or that what we can know, at least, is closed under known entailment (C). Yet, this assumption is false, Nozick maintains. Therefore, Nozick concludes, the skeptical scenarios impose no barrier on our having knowledge about the external world.

Nonetheless, we should be hesitant to deny (C). We should be hesitant because the denial of (C) leads to consequences which are far more counterintuitive than the skeptical conclusion. The following objection originated with Saul Kripke.[46] Suppose that some individual, Smith, is driving through some farmland somewhere in the Midwest. He looks out the window and, on the basis of his sensory experience, comes to believe that there is a barn in the distance. Suppose also that there is a movie being filmed in the area and that many of the supposed barns are merely props. All of the props, however, are white. But, as

---

[45] *Ibid.*, p. 201.

[46] In lectures which are not yet published. (Kirk Ludwig, personal communication)

things turn out, the barn that Smith sees is red and, pointing to the barn, Smith says "that's a red barn." Smith believes that that is a red barn In this case, Nozick's conditions are met:

(1)   It is true that that's a red barn

(2)   Smith believes that that's a red barn.

(3)   If it were not true that that is a red barn, Smith would not believe that it is a red barn.

(4)   If it were true that that is a red barn, Smith would believe that it is a red barn.

On Nozick's account, it follows that Smith knows that that is a red barn. But, notice that if what Smith had pointed to were not red, then condition (3) would not be satisfied. Thus, on Nozick's account of propositional knowledge, Smith can know that that is a *red barn*, but he cannot know that that is a *barn*. This result is so strikingly counterintuitive that we ought to regard Nozick's analysis of knowledge as absurd.

Indeed, it is because we think that what we can know is closed under known entailment that we think this result is utterly absurd. I think this tells us something important about the strength of our intuitions about the truth of assumption (C). It looks as though anyone who aims to discharge the threat of skepticism by rejecting (C) is going to have to explain away these intutions. And this, I think, is a daunting task.

Rejecting I: Goldman's Historical Reliabilism

In general, reliabilism holds that a belief is justified just in case it is the result of a reliable belief-forming process. Alvin Goldman has defended a version of reliabilism, historical reliabilism, according to which a belief is justified just in case it was formed by a

reliable cognitive process.[47] Goldman's analysis begins with the question: What is a theory

of justified belief? His answer is that it is a set of principles which give truth conditions for

all possible instances of the schema: S's belief in p at time t is justified. Goldman contends

that an illuminating account of justified belief must meet the following restriction:

> (R) An adequate account of justified belief must not contain any epistemic terms
> in the explicandum.[48]

Goldman proceeds by first discussing some shortfalls with four traditional accounts of

justified belief, all of which either violate the above restriction or are open to obvious

counter-examples. It will be useful to discuss Goldman's objections to these traditional

accounts of justified belief at length, since, as I will argue, his own analysis of justified belief

violates condition R.

> (1) If S believes p at t, and p is indubitable for S (at t), then S's belief in p at t is
> justified.

> (2) If S believes p at t, and p is self-evident, then S's belief in p at t is justified.

> (3) If p is a self-presenting proposition, and p is true for S at t, and S believes p at t,
> then S's belief in p at t is justified.

> (4) If p is an incorrigible proposition, and S believes p at t, then S's belief in p at t is
> justified.

---

[47] Alvin Goldman, "What is Justified Belief?" pp. 105-130.

[48] *Ibid.*, p. 106.

To see that (1) violates condition R, we must first understand what 'indubitable' means, for there are two senses in which a belief may be said to be indubitable. In the first sense, to say that S's belief that p is indubitable is just to say that S has no grounds for doubting p. Since 'ground' is an epistemic term, on this reading of 'indubitable,' (1) violates condition R. In the second sense, to say that S's belief is indubitable is to say that S is psychologically incapable of doubting p. Although this would not violate R, it is surely incorrect. For example, a freshman taking an introductory course in philosophy (or, perhaps, even a graduate student taking epistemology) may be incapable of doubting his or her professor; this, however, does not show that his or her belief is justified.

To evaluate (2), we need to understand the expression 'self-evident,' which admits of various readings. On one reading, to say that a proposition is self-evident is just to say that it is directly justified, which obviously violates condition R. Another reading is that to say that p is self-evident is to say that it is impossible to understand p without believing it. The difficulty here focuses on how we are to understand the intended modal force of 'impossible,' which again admits of at least two readings--humanly impossible and logically impossible. On either reading, (2) is inadequate. To take the first reading, it might be humanly impossible to doubt the proposition that some events have a cause, but it is doubtful that our inability to doubt such propositions is justification conferring. The second reading may be rejected immediately, since it would render (2) vacuous.

To evaluate (3) we need to understand what 'self-presenting' means. One suggestion, is the following: "Proposition p is self-presenting if and only if: necessarily, for any S and

any t, if p is true for S at t, then S believes p at t."[49] Once again, we need to understand the intended modal force of 'necessarily'. Goldman considers two readings: nomological necessity and logical necessity.

On the first reading, (3) is open to the following counter-example. Suppose that there is a nomological necessary connection between being in a certain neurological state, B, and believing a some particular proposition p, such that whenever S is in brain-state B, S is justified in believing that p. This cannot be correct, since it is easy to imagine cases in which S is in brain-state B, but where it is intuitively incorrect to say that S is justified in believing p. "For example, we can imagine that a brain-surgeon operating on S artificially induces brain-state B."[50] Goldman gives the following counter-example to the second reading of (3). "Let p be the proposition 'I am awake' and assume that it is logically necessary that if this proposition is true for some person S and time t, then S believes p at t. This assumption is consistent with the further assumption that S frequently believes p when it is false, e.g., when he is dreaming."[51] This shows, Goldman claims, that self-presenting propositions are not always justified.

In order to evaluate (4), we need to understand the notion of incorrigibility. A proposition is said to be incorrigible just in case: "necessarily, for any S and any t, if S believes p at t, then p is true for S at t."[52] Once again, the evaluation turns on the intended modal force of 'necessarily'--whether it is nomological necessity or logical necessity. On

---

[49] *Ibid.*, p. 110.

[50] *Ibid.*, p. 110.

[51] *Ibid.*, p. 111.

[52] *Ibid.*, p. 111.

the former reading, (4) is open to a counter-example similar to the one advanced against the nomological reading of (3). For example, suppose that there is a nomologically necessary connection between S's believing that S is in brain-state B and S's actually being in brain-state B, such that S is in brain-state B if and only if S believes he is in brain-state B. From this it follows that "I am in brain-state B" is an incorrigible proposition for S. Yet, if brain-state B were artificially induced by a neurosurgeon, S's belief that he is in brain-state B would be incorrigible, but nonetheless his belief would not be justified. Let us turn to the latter case, where the necessity in question is taken to be logical necessity. Suppose S comes to believe some complicated mathematical truth (p) on the basis of, say, tarot cards. Since all mathematical truths are logically necessary, if S believes that p, then S's belief is true, and hence p is incorrigible for S. On this reading of (4), S's belief that p would be justified--contrary to our intuitions.

Goldman's own account of justified belief is motivated, to a large extent, by the inadequacies of the traditional accounts. One primary difficulty with these approaches, Goldman claims, is that they assume "whenever a person has a justified belief, he knows that it is justified and knows what the justification is."[53] Moreover, "it is further assumed that the person can state or explain what his justification is."[54] The problem, in a nutshell, is that these traditional approaches are all internalist--they assume that for something to be justification-conferring, it must be cognitively accessible to the believer.

---

[53] *Ibid.*, p. 106.

[54] *Ibid.*, p. 106.

Goldman advocates an externalist theory of justification. On such a theory, the justification conferring property or process need not be cognitively accessible to the believer. There are two primary motivations for adopting externalism. The first is that it does seem intuitively plausible to hold that someone could have a justified belief without being able to cite reasons for holding that belief. The second is that the traditional internalist accounts of justification have not met with considerable success. As Goldman notes, "each of the foregoing attempts confers the status of 'justified' belief without restriction on *why* the belief is held, i.e., on what *causally initiates* the belief or *causally sustains* it."[55]

On Goldman's view, a belief is justified just in case it was produced by a justification-conferring, belief-forming process, where "a 'process' is a functional operation or procedure, i.e., something that generated a mapping from certain states--'inputs'--into other states--'outputs'."[56] Goldman gives the following examples of belief-forming processes which are justification conferring: standard perceptual processes, remembering, good reasoning and introspection. These processes are construed as functional operations. The outputs are beliefs and other cognitive states. The inputs of the reasoning operation are antecedent beliefs and entertained hypotheses. Another example is the memory process, "which takes as input beliefs or experiences at an earlier time and generates as output beliefs at a later time."[57]

---

[55] *Ibid.*, p. 113.

[56] *Ibid.*, p. 116.

[57] *Ibid.*, p. 117.

What these justification-conferring, belief-forming processes all have in common, Goldman claims, is that they are all reliable, that is, they tend to produce true beliefs. His claim is that:

> The justificational status of a belief is a function of the reliability of the process or processes that cause it, where . . . reliability consists in the tendency of a process to produce beliefs that are true rather than false.[58]

This is only an initial suggestion. Indeed, Goldman arrives at his final position by means of dialectic--he presents an analysis of justified belief, criticizes that analysis, and then offers a new formulation which attempts to meet that criticism.

His first suggestion goes as follows:

(5) If S's believing p at t results from a reliable cognitive belief-forming process (or set of processes), then S's belief in p at t is justified.

Now, since the notion of a reliable belief-forming process has been cashed out in terms of a tendency (a statistical propensity) to produce true beliefs, (5) meets condition R. The difficulty with (5), however, is that it does not account for the possibility that S's antecedent beliefs may be largely false. In light of this, Goldman introduces the notion of conditional reliability, which holds that "a process is conditionally reliable when a sufficient proportion of its output-beliefs are true *given that its input-beliefs are true*."[59] Goldman also distinguishes between what he calls 'belief-dependent cognitive processes' and 'belief-

---

[58] *Ibid.*, p. 115.

[59] *Ibid.*, p. 119.

independent cognitive processes'. The former type of cognitive processes are functions whose inputs include beliefs, and the latter type of cognitive process are functions whose inputs do not include beliefs. Goldman offers the following replacement for (5):

(6_A) If S's belief in p at t results ('immediately') from a belief-independent process that is (unconditionally) reliable, then S's belief in p at t is justified.

(6_B) If S's belief in p at t results ('immediately') from a belief-dependent process that is (at least) conditionally reliable, and if the beliefs (if any) on which this process operates in producing S's belief in p at t are themselves justified, then S's belief in p at t is justified.[60]

If (6_A) and (6_B) are supplemented with a standard closure clause, Goldman claims, the result is a complete theory of justified belief--a theory he dubs 'historical reliabilism'. "The theory says, in effect, that a belief is justified if and only if it is 'well-formed,' i.e., it has an ancestry of reliable and/or conditionally reliable cognitive operations."[61]

Goldman considers the following sort of objection which forces a slight modification to (6)--call this the 'case of the persistent believer.' Imagine a case in which some individual, Jones, is led to believe that his memory beliefs from a certain period of time of his life are, for the most part, mistaken--say, his memory beliefs from his two tours in Vietnam. We can imagine that Jones has it on the highest authority (U.S. government officials) that while in Vietnam he was exposed to a chemical compound which causes amnesia in 98% of the population. As things turn out, these government officials are well aware of the fact that Jones is among the 2% who are unaffected by exposure to the

---

[60] *Ibid.*, p. 119.

[61] *Ibid.*, p. 119.

compound; nonetheless, they are under orders not to inform Jones of this fact. Although Jones listens to what these officials tell him and has excellent reason to believe them, he persists in his beliefs formed on the basis of the ostensible memories from his two tours in Vietnam. Although these beliefs result from reliable belief-forming processes, we are not inclined to say that they are justified since Jones has evidence (albeit misleading evidence) that his memory beliefs were not the result of a reliable belief-forming process.

In light of counter-examples of this sort, Goldman considers a number of proposals for modifying (6_A). After rejecting a number of proposals, Goldman arrives at his final analysis of justified belief.

(10) If S's belief in p at t results from a reliable cognitive process, and there is no reliable or conditionally reliable process available to S which, had it been used by S in addition to the process actually used, would have resulted in S's not believing p at t, then S's belief in p at t is justified.

Notice that (10) does block the case of the persistent believer. There is another belief-forming process available to Jones which is such that if he had used that alternative process, in addition to the process he actually used, he would not have believed that his memory beliefs from his two tours in Vietnam were veridical. This alternative belief forming process would take into consideration the evidence that was presented to him by the government officials.

Goldman himself notes that there is a difficulty with (10). The difficulty concerns the notion of an *available* belief-forming process. Goldman himself puzzles over the question of what it means for a process to be available to a cognizer--"were scientific procedures

'available' to people who lived in pre-scientific ages?"[62] This requirement would be much too strong. Moreover, "it is implausible to say that all 'available' processes ought to be used, at least if we include such processes as gathering new evidence."[63]

So, in order for (10) to get the right results, it looks as though Goldman needs some principled way of marking off a distinction between those processes which are considered available and those that aren't. He offers the following suggestion. "What I think we should have in mind here are such additional processes as calling previously acquired evidence to mind, assessing the implications of that evidence, etc."[64] The problem with this suggestion, however, is that it makes use of the expressions 'evidence' and 'assessing the implications of that evidence'. So, in order for his analysis of justified belief to get the right results, it looks as though Goldman is forced to appeal to epistemic terms, which violates condition R. So, as stated, by his own standards, Goldman's final analysis of justified belief proves inadequate.

There is another, more serious difficulty with Goldman's final analysis. Again, Goldman himself points out the difficulty. Since Goldman is giving a conceptual analysis of justified belief, the analysis must apply to all possible cases. This suggests that "for any cognitive process C, if C is a reliable process in possible world W, then any belief in W that results from C is justified."[65] So, imagine a possible world in which wishful thinking is a reliable belief-forming process. Say, a world in which "a benevolent demon so arranges

---

[62] *Ibid.*, p. 126.

[63] *Ibid.*, p. 127.

[64] *Ibid.*, p. 127.

[65] *Ibid.*, p. 122.

things that beliefs formed on the basis of wishful thinking is reliable."[66] In such a world,

beliefs formed on the basis of wishful thinking would be justified, on Goldman's account.

Now, as one might expect, Goldman wants to rule out this case. He gives the following

suggestion for ruling out such a possibility:

> Instead of construing the theory as saying that a belief in possible world W is justified if and only if it results from a cognitive process that is reliable in W, we may construe it as saying that a belief in possible world W is justified if and only if it results from a cognitive process that is reliable in our world.[67]

Moreover, Goldman claims:

> We note certain cognitive processes in the actual world, and form beliefs about which of these are reliable. The ones we believe to be reliable are then regarded as justification-conferring processes. In reflecting on hypothetical beliefs, we deem them justified if and only if they result from processes already picked out as justification-conferring. . . . Since wishful thinking is not among these processes, a belief formed in a possible world W by wishful thinking would not be deemed justified, even if wishful thinking is reliable in W.[68]

At first blush, it looks as though this response works only if we already have some

means of picking out which belief-forming processes are reliable in the actual world. If this

is right, then Goldman would have to presuppose that the actual world in not a world in

which a benevolent demon arranges things so that wishful thinking is a reliable belief-

forming process. However, this would be tantamount to presupposing that skepticism is

false.

---

[66] *Ibid.*, p. 122.

[67] *Ibid.*, p. 123.

[68] *Ibid.*, p. 122.

Goldman would, of course, deny this. On Goldman's view, one doesn't need to be able to show that one's belief-forming mechanisms are reliable in order for them to be so and so for one to have knowledge and justified belief. If this is correct, then contrary to our initial criticism, Goldman would not have to have a proof that the actual world is one in which wishful thinking doesn't lead to true belief. Nonetheless, this response is unconvincing. In effect, the most Goldman has to say is that if historical reliabilism is true, then the skeptic's claim that knowledge is impossible has been shown to be false, since no one denies that our beliefs could be caused by reliable belief-forming mechanisms.

This is unsatisfying because an analysis of knowledge should tell us in general terms what makes for justified belief, independently of reference to the way things actually are, rather than simply providing a general characterization of what is required for knowledge and then leaving it up to the world whether anything meets the requirements. If the concept of knowledge is in that way a general concept, then whatever features are sufficient for knowledge in the actual world will be sufficient for knowledge in other possible worlds as well.

Goldman does consider the possibility that the actual world could be a possible world in which a benevolent demon arranges things so that wishful thinking is a reliable belief-forming process. The actual world might be such a world

> . . . because, unbeknownst to us at present, there is a benevolent demon who, lazy until now, will shortly start arranging things so that our wishes come true. The long-run performance of wishful thinking will be very good, and hence even the

new construal of the theory will imply that beliefs resulting from wishful thinking (in our world) are justified.[69]

Goldman's response to this objection is that it "contravenes our intuitive judgment on the matter." This response is grossly inadequate. Recall the question with which Goldman begins his investigation--namely, "What is a theory of justified belief?" His answer to this question was that a theory of justified belief is a set of principles which give truth conditions *for all possible instances* of the schema: S's belief in p at time t is justified. In light of the above counter-example, Goldman simply shifts the focus of his investigation. He says:

> What we really want is an explanation of why we count, or would count, certain beliefs as justified and others as unjustified. Such an explanation must refer to our beliefs about reliability, not to the actual facts. The reason we count beliefs as justified is that they are formed by what we believe to be reliable belief-forming processes. Our beliefs about which belief-forming processes are reliable may be erroneous, but that does not affect the adequacy of the explanation.[70]

Nonetheless, it is clear that giving such an explanation is an entirely different project than giving truth conditions for all possible instances of the schema: S's belief in p at time t is justified. Moreover, if his goal is to give an explanation (in the sense described above), then his analysis of justified belief offers little promise of resolving the skeptical challenge.

---

[69] *Ibid.*, p. 123.

[70] *Ibid.*, pp. 123-24.

Rejecting E: Armstrong's Direct Realism.

If we have non-inferential knowledge of the world around us, as the Direct Realist claims, then it's false that all we start off knowing is the contents of our own thoughts. What exactly is "non-inferential" knowledge and why should we think that we do indeed know things non-inferentially? David Armstrong has addressed both of these questions. He gives the following analysis of non-inferential knowledge.

A knows p non-inferentially if, and only if, A has no good reasons for p but:

(i)   A believes p;

(ii) p is true;

(iii) A's belief-that-p is empirically sufficient for the truth of p.

(Since if x is necessary for y then y is sufficient for x, condition (iii) can also be written:

(iii) The truth of p is empirically necessary for A's belief that p.[71]

So, for example, I know non-inferentially that I am currently seated in front of my computer typing away iff I truly believe that I am seated in front of my computer, and I could not have empirically had the belief that I am seated in front of my computer if it were not the case that I am currently seated in front of my computer.

One point to note is that if conditions (i) and (iii) are satisfied then condition (ii) will be satisfied trivially. In light of this, Armstrong simplifies his account of non-inferential knowledge as follows.

---

[71] David Armstrong, *A Materialist Theory of Mind,* p. 189.

A knows p non-inferentially if, and only if, A has no good reasons for p but:

 (i) A believes p;

 (ii) A's belief-that-p is empirically sufficient for the truth of p.

It is clear that Armstrong's analysis of non-inferential knowledge hinges crucially on the tenability of the notion of empirical sufficiency.

  So, what is it for someone's belief to be empirically sufficient for the truth of some proposition? Although Armstrong offers the following suggestion, he is content, on the final analysis, to leave unexplicated the notion of empirical sufficiency. The most he says is the following:

> I would argue that it [something's being empirically sufficient for something else] is a matter of being able to subsume the situation under a covering law of a certain form, a law which, however, may be quite unknown to us even though we are certain that the situation before us is a case where something is empirically sufficient for something else.[72]

Nonetheless, immediately after offering this suggestion on how we should understand the notion of empirically sufficiency, he remarks:

> I deliberately refrain from making this part of the analysis of what it is to know something non-inferentially. The notion of something being empirically sufficient for something else in a certain situation is, I think, an intuitively clear one, and we expose the minimum of flank if the analysis is rested simply on this notion.[73]

---

[72] *Ibid.*, p. 190.

[73] *Ibid.*, p. 190.

So, Armstrong proposes to rest his account of non-inferential knowledge on the unexplicated notion of empirical sufficiency. We should, therefore, conclude that his analysis of non-inferential knowledge is, at best, woefully incomplete.

We have, thus far, addressed only the first of our two-part question. Let us turn, now, to the second part of our initial question--why should we think that we have non-inferential knowledge? Armstrong's answer is that if we are to have any knowledge at all, we must have non-inferential knowledge. His argument is found in the following passage.

> To say I have good reasons for a proposition, p, is to say that I know other truths, q, which constitute good reasons for p. (If I did not *know* q, my reasons would not be good.) This new knowledge will itself require the possession of further good reasons, and so to infinity. The vicious infinite regress exposes the concealed circularity in the proposed definition of knowledge: knowledge is true belief, together with knowledge of good reasons for that belief.
> This difficulty makes it clear that, if knowledge is a meaningful notion at all (and I propose to assume that it is), all knowledge based on good reasons must terminate in knowledge that is not based upon any reasons at all [i.e., non-inferential knowledge].[74]

This argument is seriously flawed. What is at question here is not whether the notion of knowledge is coherent. Everyone assumes that it is. What is at question is whether we possess knowledge, non-inferential or otherwise. So, if Armstrong's discussion is to be considered even so much as germane to the issue at hand, then we have to take him as affirming the following conditional: "if we are to have any knowledge at all, then our knowledge must terminate in non-inferential knowledge." Yet Armstrong leaves the notion of non-inferential knowledge at an intuitive level. For this response to come off, Armstrong

---

[74] *Ibid.*, p. 188.

needs to explicate this notion. He needs to provide criteria for distinguishing inferential from non-inferential knowledge. I conclude, then, that Armstrong's response is, at best, incomplete.

In considering these responses to the skeptical challenge, we are not only aiming to motivate an examination of rejecting (M), the claim that there is a necessary connection between the contents of our thoughts and the actual state of our environment, we are also aiming to discover what mistakes we need to avoid, if we are to find an adequate solution to the problem of skepticism. This issue will be examined more fully in the next section of this chapter, where we shall set down some success conditions on responses to the problem of skepticism about the external world.

Rejecting M: Berkeley's Idealism

George Berkeley, the first philosopher to advocate idealism in its modern guise, held that the problem with the metaphysical assumption turns on just how we are to understand the term 'external world'. He argues that the notion of *external world* that is employed in the above argument is incoherent. He claims that the metaphysical independence assumption makes sense only on a representational theory of perception, and that this theory is incorrect. The representational theory of perception, in a nutshell, holds that our ideas (the content of our perceptual beliefs) represent things in the external world (material objects), where these material objects are thought to be the cause of what we are immediately aware of in perception. Representative realism holds that our perceptual beliefs are veridical just in case they were caused by the (independently existing) material objects they represent.

Berkeley has three principal objections to this view. The first is that only an idea can resemble another idea, so it is incoherent to hold that ideas resemble independently existing material objects. Hence, the notion of an independently existing external world is incoherent. The second objection is that the above view could be correct only if it were possible for external objects to exist while not being perceived. This, however, is absurd, Berkeley claims, since objects can exist while not being perceived only if it is possible to conceive of an object which is not perceived. But, simply in virtue of conceiving an object, that is, by calling it before one's mind, it follows that one is perceiving it. Hence, it is not possible for objects to exist while not being perceived. The third objection focuses on the intelligibility of the notion of substance. On the above view, substance is conceived of as a substratum which 'supports' various qualities or properties. However, it is impossible to conceive of a bare particular, that is, something which has no properties. So, it looks as though we have to take the term 'support' metaphorically. If so, we have (and can have) no understanding of the notion of substance. The notion of a material substance, Berkeley concludes, is nothing but a fiction invented by philosophers.

How, then, are we to understand the term 'external world'? On Berkeley's view, the external world is just the common sense world that we talk about in ordinary, everyday life. When the ordinary person talks about tables and chairs, rocks and trees, he is talking about the ordinary objects of perception. Now, since all we ever directly perceive are ideas, and what we directly perceive are tables and chairs, rocks and trees, and the like, it follows that such objects are simply ideas (or collections of ideas). In short, the external world is mind-dependent.

Berkeley's own version of idealism, according to which ordinary objects are simply collections of ideas, has given way to more sophisticated versions of the same idea. One difficulty with Berkeley's view is that he is working with a mistaken conception of the nature of mental states and consciousness. He thought, incorrectly, that every conscious state was a matter of a thinking being in some relation to an idea. Today most philosophers reject this analysis of consciousness in favor of taking consciousness to involve simply awareness of the states of the self.

Berkeley also failed to draw a number of crucial distinctions in discussing the contents of the mind. Of particular significance is his conflation of modifications of consciousness and propositional attitudes, and their contents. Following in Locke's footsteps, Berkeley used a single world, 'idea,' to cover everything which might be thought to be the objects of the mind's attention. This obscures a number of important distinctions, and undermines a number of Berkeley's arguments which rest on the failure to draw these distinctions.[75]

Nonetheless, the core idea of Berkeley's response to skepticism can be preserved in a more adequate framework for talking about the mind. The key element in Berkeley's strategy is to interpret talk about the objects of common sense beliefs to be really about things or logical constructions of things of which we are immediately aware. This approach gives rise to the version of idealism known as linguistic phenomenalism, according to which all talk about ordinary objects of common sense knowledge, such as rocks and trees, can be

---

[75] I will not here rehearse these well-known difficulties. For a sympathetic, but critical account, see A.C. Grayling, *Berkeley: The Central Arguments.*

exhibited to be synonymous with sentences in a language which talks only about what one is immediately aware of in consciousness and what one would be aware of in various circumstances.

It is not my aim here to provide an exhaustive critique of linguistic phenomenalism, or all the varieties of this idea. Nonetheless, it will be appropriate to indicate what I see as the chief obstacles to this view, since I wish to motivate the examination of externalist theories of thought content as the only remaining live option for dealing with the problem of skepticism about the external world.

I wish to say first off that I regard the chief difficulty for this view to be its prima facie implausibility. For instance, when I talk about plant fossils which predate the existence of the human race, it is obvious that I am talking about something which existed before the existence of any thinking beings. When I talk of nebulae and galaxies, I am quite sure that I intend to be taking about things which will in all likelihood outlast the sorts of thinking beings that currently exist, and which certainly could have existed independently of any thinking beings. I shall call this the "objection from modal intuitions." I think it is one of the strongest objections there is against the view that the world is mind-dependent. A proponent of linguistic phenomenalism must not only show that he can provide a plausible structural "translation" of talk of the spatial world into talk of an individual's experience, on which more in a moment, but must also explain away our modal intuitions which are incompatible with his thesis.

In addition to this difficulty, there is the difficulty that linguistic phenomenalists have simply not managed to provide any plausible translation schemes for representing talk of the

spatial world in terms of talk of experiences. In particular, there are two difficulties which I think are most pressing for the linguistic phenomenalist. Let us consider an example. Suppose that we want to translate the sentence "There is a six-pack of beer in my refrigerator." We cannot translate this into a statement about the current experiences I am having, because the statement may be true even though I am not in my kitchen. So I have to appeal to what I would experience in certain circumstances. And even if I were in my kitchen, with my refrigerator door open, looking inside, the current experience would not be sufficient for the truth of the statement because we would need to rule out various ways in which the appearance that there is a six-pack of beer in my fridge is misleading. What we want is a translation schema of the following sort:

> There is a six-pack of beer in my fridge iff$_{df}$ if I am in my kitchen, positioned in front of my fridge, with the door open, then I am having experiences as of a six-pack of beer sitting on a shelf and if I were to reach into the fridge, and pull out a chilled beer, I would feel a chilled object . . . ., or, if I am not in my kitchen, if I were to go into my kitchen, position myself in front of my fridge, open the door, then. . . .

Now we can begin to see the enormous difficulties that plague this approach. First, the descriptions of the sorts of experiences we are supposed to get have been couched in terms of the sorts of objects they would represent, but this is to use the vocabulary that we originally wanted to show how to eliminate. So, one difficulty with this approach, one which has yet to be resolved, is to provide a kind of description of our experiences which does not build in representational content that is of the kinds of objects we want to reduce concepts of pure qualitative experience types. This looks hopeless since the qualitative content of

experiences is infected, as it were, with representational content (consider for a simple example how a Necker cube looks in each of the ways it can appear to one--that difference is representational in character.)

Second, even apart from this, it is clear that the right-hand side of our biconditional includes a lot of physical objects and spatial world vocabulary in the characterization of the counterfactual circumstances in which we would expect to get various experiences! We can try to eliminate that, but the trouble is that the same difficulty emerges for each of the terms. And it is not accidental. It is because we do conceive of the world as independent of us and the cause of our experiences, so of course we describe the circumstances in which we would expect to get experiences appropriate for veridical experiences of physical objects by describing the physical circumstances in which we would expect the usual causal conditions sufficient for this to obtain. This is another manifestation of the fact which gives rise to the objection from modal intuitions. We really do conceive of the physical world as conceptually independent of the existence and nature of any thinking being.

The second main difficulty I want to point out is also connected with this. That is that it is hard to see that the end point of linguistic phenomenalism leaves any room for the thought that any being exists except the experiencing self (for each of us, his or her own self). To see this, consider how we would make sense of talk of others. This too, since they are thought of as having a place in the world of physical things, would get reduced to a "permanent possibility of sensation," in Mill's handy phrase. But then since the experiencing self is not so reduced, it turns out that we are no longer in a position to think of other thinking beings as thinking beings in the same sense that we are ourselves. And

thus, we are ultimately reduced to logical solipsism, the view that only we ourselves exist. I submit that this is a fate far worse to suffer and more repugnant to common sense than even the most radical forms of skepticism about the possibility of justified belief about the external world.

V. Success Conditions on Responses to Skepticism

Thus far, we have examined a number of traditional responses to the skeptical challenge and we have found them inadequate. Our aim in this section is to use the analysis of these inadequacies to set down some success conditions on responses to the skeptical challenge. Two points are especially important. The first is that an adequate response to the skeptical challenge must not appeal, at any point, to empirical considerations. The second point is that an adequate response must identify the mistake in the skeptic's reasoning. Let us examine these points in more detail.

Goldman's response, based on the rejection of (I), tacitly appeals to our ability of distinguish reliable from non-reliable belief forming processes. Armstrong's analysis of non-inferential knowledge proceeds from the assumption that we can and do possess empirical knowledge. Quine, himself, acknowledges that the appeal to empirical considerations is circular; and, as Stroud has noted, if this doesn't undermine Quine's project, that is only because Quine is engaged in a project different from that of the traditional epistemologist. The first success condition that I shall insist upon is that an adequate response to the skeptical challenge must not appeal to propositions which fall within the domain of the skeptic's conclusion.

The second success condition that I want to insist upon is that an adequate response must provide a diagnosis of the skeptic's mistake. Naturalized epistemology fails most clearly on this score. It fails because it does not explain away the intuitive appeal of the Cartesian thought experiments. What we want out of a good response is an analysis of what has gone wrong. In other words, an adequate response must provide an alternative interpretation of the sorts of thought experiments on which philosophical skepticism rests.

There are some further success conditions that I want to mention briefly. Although we are focusing on skepticism about the external world, it should be clear from my analysis of the assumptions underling skepticism that an adequate response ought to be generalizable in the sense that a good response to the problem of skepticism about the external world ought to provide a template for resolving the skeptical challenge with respect to other domains. Recall that I argued in section III of this chapter that skepticism about other minds and skepticism about induction rest on the same assumptions as skepticism about the external world. If this is right, then if we can solve the problem of skepticism about the external world, then we ought to be able to resolve the problem of skepticism about other minds and about induction.

VI: The Externalist Response

Like Idealism, Content Externalism aims to respond to the skeptic by denying (M), the assumption that there is no necessary connection between the contents of our thoughts and the actual state of our environment. But instead of making the world depend on the

mind, the externalist makes the mind depend on the world. In a nutshell, externalism is the view that the identity conditions for the contents of our propositional attitudes include facts about our actual environment and our causal interaction with that environment.

Over the next three chapters we shall examine the theories of three prominent externalists: Tyler Burge, Hilary Putnam and Donald Davidson. I have chosen to discuss the views of these particular philosophers because of the clarity, originality, and depth of their thinking. Nonetheless, I think the arguments these philosophers advance are ultimately unsuccessful. In what follows, I aim to show that (a) externalism is false and (b) even if externalism were true, it would not provide an adequate response to the skeptical challenge.

CHAPTER III

BURGE'S ARGUMENTS FROM ANTI-INDIVIDUALISM

<u>Introduction</u>

In the previous chapter, I argued that the Cartesian thought experiments were intended to motivate a certain picture of the mind-world relation, one on which the mind is construed as being wholly independent of the external world--independent in the sense that propositions about the mind neither entail, nor are entailed by, propositions about the external world. Without this assumption, the argument for skepticism collapses.

In "Cartesian Error and the Objectivity of Perception" Tyler Burge argues that the Cartesian thought experiments do not provide any straightforward support for the conception the mind-world relation needed by the skeptic--a view which Burge labels "individualism."[1] In addition, he advances a positive argument based on considerations about perception which suggests that this conception of the mind-world relation is misguided.[2]

The aim of this chapter is first to clarify the target of Burge's attack, and then to determine whether his attack is successful. I begin with an examination of Burge's landmark essay "Individualism and the Mental." In this article, Burge argues that the identity

---

[1]   Tyler Burge, "Cartesian Error and the Objectivity of Perception," p. 68.

Now, it should be noted at the outset that Burge is not concerned (at least in the article under discussion) with the problem of skepticism about the external world. Nonetheless, his discussion is relevant to the issue of skepticism insofar as some form of "individualism" is required for the skeptic's argument to go through.

conditions for the contents of our propositional attitudes include facts about our social environment--the linguistic conventions of our community. I then turn to an examination of the anti-skeptical argument Burge advances in "Cartesian Error and the Objectivity of Perception."

This chapter is divided into six sections. In the next section, I discuss Burge's famous "Twin Earth" thought experiment (not to be confused with Putnam's, the subject of the next chapter), the central aim of which is to show that the identity conditions for the contents of propositional attitudes include facts about the linguistic conventions of the community of which one is a member. Section III contains an analysis of Burge's defense of social externalism. In section IV, I argue that social externalism is inherently flawed. The thrust of my criticism is that Burge's account of how beliefs are individuated leads to a contradiction. I conclude that social externalism is false. In addition, I argue that even if social externalism were true, the most one could establish on the basis of it is a claim of a conditional form: that if we live in a non-solipsistic world, then linguistic conventions may be relevant to the determination of thought content. However, to affirm the antecedent of this conditional would be to beg the question against the skeptic.

In section V, I discuss the anti-skeptical argument Burge advances in "Cartesian Error and the Objectivity of Perception." His argument proceeds in two stages. In the first stage, he argues that the Cartesian thought experiments fail to establish what he calls the "individualistic assumption." In the second stage, he argues that without this assumption, there is no skeptical conclusion to be drawn. However, if my analysis of the assumptions underlying philosophical skepticism is correct, then the skeptic is not committed to the

individualist assumption, but only to a much weaker assumption. This weaker assumption survives Burge's attack. I conclude that Burge's anti-skeptical argument is unsuccessful.

In section VI, I turn my attention to Burge's Argument for Perceptual Externalism. The aim of the argument is to show that the contents of our perceptual experiences are individuated in part by relations with our physical environment. In section VII, I argue that the most that we can get from perceptual externalism is a conclusion of conditional form: if we live in a world that admits of such relations, then facts about our environment are relevant to the determination of perceptual content. Yet to affirm the antecedent of this conditional is to beg the question against the skeptic.

II. Burge's argument from "Twin Earth"

In "Individualism and the Mental" Tyler Burge considers a number of related thought experiments, all of which are designed to show that the identity conditions for the contents of propositional attitudes include facts about the social environment.[3] The sorts of propositional attitudes to which Burge's thought experiments appeal are ordinary beliefs and desires, such as believing that one has arthritis in one's knee or desiring pork brisket for dinner, and what are, Burge claims, beliefs and desires involving slight misconceptions of concepts expressed in such attitudes, for example, believing that one has arthritis in one's *thigh*, or desiring *clam* brisket for dinner. The identity conditions for the contents of proposi-

---

[3] Tyler Burge (1979). Unless otherwise stated, all parenthetical citations within the text are from *The Nature of Mind*, D. Rosenthal, (ed.), Oxford: Oxford University Press, 1991, pp. 536-567.

tional attitudes are those conditions which must obtain in order for an individual to have attitudes with the specified contents.

Burge's conclusion rests on the following three-step thought experiment. To begin, Burge asks us to imagine a case of incomplete understanding in which an individual misconstrues (incompletely or partially understands) some notion putatively involved in the contents of some of his thoughts (step 1).[4] In the next step (step 2), we consider a counterfactual supposition. We hold the actual individual's life history (asocially, non-relationally and non-intentionally described) and physiology constant, and suppose that the linguistic practices of the counter-factual community are such that the individual's actual incomplete understanding of the particular notion now reflects complete understanding, as determined by his (counterfactual) linguistic community (i.e., his use of the relevant term accords with the counterfactual community's linguistic conventions).

The next step (step 3) involves an interpretation of the counterfactual case. Actually and counterfactually described, the individual has the same dispositions to affirm (or deny) sentences which contain the particular expression. Even so, the actual individual's verbal behavior reflects an incomplete understanding of the relevant notion, whereas the

---

[4]   Burge's exposition of the thought experiments makes use of the expression "understanding (mastering) a notion." When he discusses cases of incomplete understanding, he does so with respect to notions, rather than talking about incompletely understanding the term that is used to express the particular notion. Burge's motivation for this way of talking lies in the observation that one could master a notion without mastering the term which is used to pick out that notion, in particular, if one speaks another language which has a synonymous term.

This is actually quite problematic in that even in Burge's thought experiments we infer that a subject incompletely understands some notion by observing that the subject uses the relevant term incorrectly, i.e., he uses the term in a manner which does not accord with the linguistic conventions of *his* community.

counterfactual individual's verbal behavior reflects a complete understanding, relative to how the particular expressions are understood in their respective linguistic communities. This follows from the stipulation that the extensions of the terms (used to express the relevant notions) differ relative to the actual and counterfactual linguistic communities.

So, for example, we are asked to imagine an individual who has incompletely mastered some notion, but whose incomplete understanding is not such as to deter ordinary attributions of the attitude. We can suppose that, as we would say, the individual believes, falsely, that brisket is a certain cut of beef, and not, say, any cut of pork.[5] Although this individual's brisket-notion isn't quite on the mark, no one at the local supermarket is going to deny that he has brisket-beliefs and that these beliefs really are about brisket. Furthermore, he could have various brisket-beliefs which many people in his linguistic community would be inclined to agree with, say, that there is no such thing as clam brisket. But still, Burge insists, the individual's belief that brisket cannot be a cut of pork is false.

---

[5]   Burge lays out his thought experiment with a different example--one involving a misconception of arthritis. In the first step of the original version of the thought experiment the individual believes, falsely, that he has arthritis in his thigh. Then we are asked to suppose that in the counterfactual case, "arthritis" refers to various rheumatoid ailments, and, hence, the individual's belief is true. But when Burge runs the reversal of the thought experiment, where it is stipulated that the individual has complete understanding of the relevant notions, the individual's belief is that he has arthritis.

This is not a true reversal of the first thought experiment; Burge claims that the only difference between the two thought experiments involves the order in which the first two steps are taken. It is, however, clear that the beliefs (and belief sentences) examined in the two thought experiments are different. The first thought experiment involves the belief that one has arthritis in one's thigh, and the second involves the belief that one has arthritis.

For this reason I have chosen to lay out Burge's argument so that the misconception involves the notion of brisket. This is an example Burge mentions (and he thinks it supports the thought experiment), but one which he doesn't develop. At any rate, the example discussed remains true to the spirit of Burge's argument and illustrates a true reversal.

Next, we imagine a counterfactual situation in which the individual's physiological makeup and life history remain constant, but in which the linguistic community uses the word "brisket" to refer only to a certain cut of beef and not to any cut of pork. The linguistic practices of the counterfactual community are such that they capture our individual's actual use of the term "brisket." And the belief of the individual in the counterfactual community that he would express by saying "brisket isn't any cut of pork" would be true.

In both the actual and counterfactual case, the individual has the same physiological states and the same life history (asocially, non-relationally, and non-intentionally described). The individual develops the disposition to assent to the sentence "brisket isn't a cut of pork" from the same proximate stimuli. The word "brisket," however, differs in extension in the two communities. In actual fact, "brisket" is used to refer both to certain cuts of beef and to certain cuts of pork, but in the counterfactual community, "brisket" is used to refer only to certain cuts of beef.

The individual's actual belief would differ from his counterfactual belief in virtue of the fact that in the two cases the subject's beliefs would be about extensionally distinct things. This difference, Burge claims, stems from social factors that are independent of the individual (asocially, non-relationally, and non-intentionally described). The individual has the same dispositions, the same physiological states, the same life experiences (described in isolation from his environment) in the actual situation as he does in the counterfactual situation; skin-inward, everything about the individual remains constant, yet the contents of his attitudes differ. The counterfactual individual would lack beliefs involving the concept of brisket; his beliefs would be said to involve the concept of brisket$_{cf}$.

How, then, Burge asks, are we to account for this difference? The only difference between the actual and counterfactual situations involves the extension (and use) of the word "brisket." Taken in isolation from the linguistic community, there is no way to distinguish (or to individuate) the actual individual's belief and the counterfactual individual's belief. Yet, we seem to be committed to the claim that the beliefs are different, simply in virtue of the fact that the beliefs are about different things--brisket and brisket$_{cf}$.

Burge's argument begins with the widely-held assumption that content clauses do not freely admit substitution of co-referring or co-extensional expressions without the possibility of changing the truth value of the containing sentence. And, as Frege pointed out, this is because co-extensional expressions within content clauses may serve to indicate cognitively different notions.

Burge's discussion of this point is mostly heuristic. His claim is this: Content clauses of propositional attitude ascriptions have traditionally been taken as a primary means of identifying a subject's intentional mental states. The motivation for this assumption is that we cannot, in general, substitute co-referring or co-extensional expressions within embedded content clauses so as to preserve the truth value of the containing sentence. Burge's line of reasoning exploits this assumption. Surely, he says, if ever co-referring expressions in oblique position can indicate different thoughts, then it is simply undeniable that obliquely occurring expressions that are not extensionally equivalent indicate different thoughts. Burge says:

> It is normal to suppose that those content clauses correctly ascribable to a person that are not in general intersubstitutable *salva veritate*--and certainly

those that involve extensionally nonequivalent counterpart expressions--
identify different mental states or events.[6]

For example, suppose the sentence "Smith believes that Rome is the most beautiful city in Europe" is true. Although "Rome" and "the capital of Italy" are extensionally equivalent, substitution within content clauses is not generally truth preserving. The sentence "Smith believes that the capital of Italy is the most beautiful city in Europe" may not correctly capture the content of Smith's belief, particularly if Smith is unaware that Rome is the capital of Italy.

This claim figures in Burge's thought experiments in that content clauses that are taken to give the attitudes of the individual, actually and counterfactually described, contain obliquely occurring expressions that are non-co-extensional in the languages in the respective communities. Burge puts this point as follows:

> On any systematic theory, differences in the extension--the actual denotation, referent, or application--of counterpart expressions in that-clauses will be semantically represented, and will, in our terms, make for differences in content.[7]

On Burge's view, extensionally nonequivalent component parts of obliquely occurring content clauses clearly call for attribution of different attitudes.

Burge's argument can be stated as follows.

---

[6] Tyler Burge, "Individualism and the Mental," p. 538.

[7] *Ibid.*, p. 537.

(1) The propositional attitudes of the actual individual that are attributed using ascriptions whose content clauses contain the word "brisket" are propositional attitudes about brisket--i.e., the actual individual's beliefs involve the concept of brisket.

(2) Asocially, non-relationally and non-intentionally described, the counter-factual individual and the actual individual have exactly the same physiological states and the same life experiences (described in isolation from their social environments), and their use of and dispositions to use the word "brisket" are also the same.

(3) The word "brisket" differs in extension in the respective linguistic communities. According to the linguistic conventions of the counterfactual community, the word "brisket" does not extend to any cut of pork, whereas in the actual community "brisket" has an extension which includes certain cuts of pork.

(4) The propositional attitudes of the counterfactual individual which are attributed using sentences whose content clauses contains the word "brisket" are *not* propositional attitudes about brisket--i.e., the counter-factual individual's beliefs do not involve the concept of brisket.

(5) Since a purely relational difference between the actual and the counter-factual situations calls for the attribution of different attitudes, the identity conditions for the contents of propositional attitudes include facts about the social environment.

According to Burge, even though the individual misconstrues the word "brisket" (and has an imprecise conception of brisket), the individual really does express beliefs about brisket when he uses the word "brisket." In fact, on Burge's view, he not only expresses beliefs about brisket, he actually *has* beliefs involving the concept of brisket. So, the argument goes, it would be correct to use the same sentence to attribute attitudes to the individual, in both the actual and the counterfactual situation, since the counterfactual individual, as we have supposed, has a complete understanding of the term "brisket," relative to his linguistic community. But, the beliefs the counterfactual individual expresses when he uses the word "brisket" are not propositional attitudes about brisket, relative to the actual

linguistic community, since in that linguistic community the word "brisket" has a different extension. The actual individual's belief and the counterfactual individual's belief are different (since they are about extensionally distinct things). The actual individual's beliefs involve the concept of brisket, whereas the counterfactual individuals' beliefs would involve the concept of brisket$_{cf}$. And the only way we can account for this difference is by appealing to social factors.

The first premise of Burge's argument relies on the claim that we could correctly attribute an attitude to an individual using a sentence whose content clause contains the term "brisket" (in oblique position), in spite of the fact that the individual--by stipulation--has an incomplete understanding of the concept of brisket. On Burge's view, even though the individual only incompletely understands the concept of brisket, it is still proper to say that he possesses the concept of brisket. This claim is based on two principles.

(P1) In cases in which a speaker only incompletely understands some concept and makes use of a term which expresses that concept, it is correct to interpret the sentence that the speaker utters (containing that term) in a straightforward and literal way, provided that his misuse of the term is not too deviant, relative to the conventions of his linguistic community.

(P2) In such cases, the content of the individual's belief is given by the proposition expressed by the sentence that the individual uses to express his belief, relative to the linguistic conventions of his community.

One consequence of these principles is that a speaker can be said to possess a concept just in case his use (and dispositions to use) a term which expresses that concept are not too deviant, relative to the linguistic conventions of his community, so as to force reinterpretation of the sentences he utters (or would be disposed to utter) which contain that term.

The difficulty here lies in figuring out just what it means to say that someone incompletely understands a concept. When we talk about concepts we typically say that someone either grasps (or fails to grasp) a concept. And, when we talk about incomplete understanding, we typically say that someone misunderstands the word used to express that concept, not the concept itself. It is natural to talk of people misunderstanding a word, but it is a bit odd to talk about people misunderstanding a concept. For example, if a person doesn't quite grasp the concept *hammer*, thinking that "hammer" extends to meat mallets, then when he utters the sentence "This is a hammer" (while pointing to a meat mallet), we would say that he isn't working with the concept *hammer*, but rather, with a concept of a hand-held, pounding tool ('pounder', for short).

Burge's contention, however, is that when, for example, someone (sincerely) utters the sentence "I fear that the brisket will be a bit tough tonight, since I didn't have a hammer handy to tenderize it," that the person is working with the concept *hammer* and not with the concept *meat mallet*. On Burge's view, since the individual's understanding (of the concept *hammer*) is not too deviant, we would typically interpret the sentence he uses literally, and attribute to him a belief involving the notion *hammer*, and not a belief involving the notion *pounder*. Burge's motivation for this claim is that in order to show that the individual's belief does not involve the concept *hammer*, it is necessary to reinterpret the sentence that the speaker uses to express his belief. Yet, Burge argues, the appeal to reinterpretation is not supported by the ordinary practice of mentalistic attributions.

According to the first premise of Burge's argument, it can be correct to attribute attitudes to individuals using terms within the content clause (of the ascription) which

express concepts that the individual only understands incompletely. As noted previously, Burge's principles of belief attribution (P1 and P2 above) depend essentially on this premise. These principles--and, thus, the first premise of Burge's argument--will be the target of my objections in what follows. Before going into that, however, it will be useful to take a look at Burge's defense of these principles.

Burge's thought experiments, as we have already seen, rest on the claim that the sentence would be used, literally and correctly, both in the actual and in the counterfactual situation to attribute attitudes with the specified contents. This is so, Burge claims, even though, as we would say, one of the subjects only incompletely understands the concept expressed by the term he uses, and the other understands the concept completely--relative to the linguistic conventions of their respective linguistic communities. So, for example, it is urged that the same sentence "So-and-so believes that he had clam brisket for dinner last night" would be used to attribute attitudes to the individual in both the actual and in the counterfactual situation, even though, it is supposed, one of the individuals uses the term "brisket" incorrectly and the other uses the term "brisket" correctly, relative to the linguistic conventions of their respective communities.

III: Burge's Defense of his Argument

The sorts of responses that Burge considers in defending his thought experiments take issue with this claim. Specifically, Burge considers two general strategies for reinterpreting

the thought experiments. The first strategy he considers for reinterpreting the thought experiments is the attempt to motivate a non-literal reading of the sentence that the individual uses to express his belief (which directly displays the subject's incomplete understanding). The idea behind this strategy is to show that if the individual doesn't fully understand the words he uses to express his belief, then those words should not be understood literally in characterizing the contents of his beliefs (relative to his language). Hence, our interpretation of the sentence used to ascribe belief contents would have to be sensitive to the subject's idiosyncratic use of the incompletely understood term.

The problem this presents for Burge's argument is that if the sentence the subject uses contains words that we know he doesn't fully understand, then that sentence should not be understood literally. If this were the case, it would not be correct to say that the subject's belief (which directly displays the incomplete understanding) is false. And, recall, Burge's grounds for distinguishing the actual individual's belief from the counterfactual individual's belief is that they differ in truth value. The move here is to say that it is not correct to say that the beliefs differ in truth value, since the sentence used by the subject containing words that he only incompletely understands needs to be reinterpreted. Hence, the same sentence, literally interpreted, could not be used to capture the contents of the beliefs of the individual in both the actual and counterfactual communities.

The second general strategy for reinterpreting the thought experiments that Burge considers attempts to sever the connection between the contents of the subject's attitudes and

the proposition expressed by the sentences which are used to attribute the contents. The second assumption that Burge needs in order to motivate the first premise of his argument, recall, holds that the contents of beliefs (and other propositional attitudes) are given by the proposition expressed by the sentence that the individual uses to express his belief. Burge sketches three arguments which seek to deny his conclusion by showing that in cases where the individual uses a term incorrectly, the proposition expressed (in his community's language) by the sentence that the individual uses couldn't possibly capture the content of his belief.

Burge notes that in common practice we routinely and ordinarily attribute attitudes to individuals using sentences whose content clause contains words that the individual only incompletely understands, even in cases in which the sentence that the individual uses to express his belief directly displays his incomplete understanding. This is simply an empirical claim. A moment's reflection on common practice renders its truth obvious. The four methods of reinterpretation that Burge examines do not deny the truth of this claim. Rather, these methods of reinterpretation downplay the significance (and reliability) of the ordinary practice of mentalistic attributions for purposes of constructing an adequate account of what we are doing when we attribute attitudes in cases of incomplete understanding. These methods are supposed to provide an alternative interpretation of the thought experiments.

The first method of reinterpreting the thought experiment that Burge considers involves an appeal to *de re* beliefs. On Burge's view, a *de re* belief is a belief which relates an individual to an actual object.[9] For example, we may characterize Alfred's *de re* belief about

---

[9] Tyler Burge, "Belief *De Re*."

some particular apple by saying "Alfred believes of the apple on the table that it is wholesome."[9] One might think, in this case, that the correct characterization of Alfred's belief depends on relational facts, since it seems that we could change the truth value of Alfred's belief by changing his environment--we could surreptitiously exchange the wholesome apple for another apple (one that is perceptually indistinguishable to Alfred) which, say, has a rotten core. Although we may change our evaluation of Alfred's belief by exchanging the wholesome apple for a rotten one, the exchange would not affect what Alfred believes *de dicto* (the actual content of his belief); we would simply change what Alfred's belief content was *of*--and, perhaps also our evaluation of whether his belief was true of it.

The appeal to *de re* beliefs could present a problem for Burge's argument if it could be shown that the cases of incomplete understanding which give rise to the thought experiments rely essentially on *de re* beliefs and not, as Burge claims, on *de dicto* beliefs. Burge describes the *de re* method of reinterpretation as follows: "Sometimes relevant mentalistic ascriptions are reinterpreted as attributions of *de re* attitudes *of* entities not denoted by the misconstrued expressions."[10]

So, to return to our example involving clam brisket (in which the actual individual's verbal behavior reflects an incomplete understanding of "brisket" and the counter-factual individual's behavior doesn't--relative to their respective linguistic communities), the appeal to *de re* beliefs would make the following suggestion: the belief that the actual individual

---

[9]   Tyler Burge, "Individualism and Psychology."

[10]   *Ibid.*, p. 548.

expresses by saying "brisket is a clam part" should be interpreted as a belief, *of* some non-brisket substance, that it is a clam part. Or, in less convoluted terms, we might say that the subject's belief was really about *that stuff* that he had for dinner last night, or *that stuff* he bought at the butcher shop.

On this method, although we would use the sentence "So-and-so believes that brisket is a clam part" to attribute a belief to the (actual) individual, the word "brisket" would be understood in a non-literal manner. We would, perhaps, following this method, put the word "brisket" in scare quotes in order to signify the non-literal reading of the term.

This line of reasoning, however, seems unpromising in that it fails to acknowledge the oblique occurrence of "brisket" within the scope of the that-clause of the sentence used to attribute the belief. This is supported by the fact that we cannot freely substitute co-referring or co-extensive expressions within the content clause of the ascription without changing our judgement that our ascription is correct.

The second method of reinterpreting the thought experiment, which Burge discusses only briefly, holds that in cases of incomplete understanding, the content of the individual's attitude is indefinite. Burge does allow that when an individual appears utterly confused, the content of the belief may indeed by indefinite. But, conflating cases of utter confusion with cases of incomplete understanding, Burge thinks, is a gross mistake. We commonly distinguish between the two sorts of cases in ordinary, everyday practice. The appeal to indefiniteness, Burge claims, is plausible only in light of the prior assumption that literal attribution requires complete understanding--the very claim that Burge's thought experiments deny.

The two remaining methods of reinterpreting the thought experiment, which Burge addresses more fully, are examined conjointly. Burge describes the first of these remaining methods as follows:

> One is to attribute a notion that just captures the misconception, thus replacing contents that are apparently false on account of the misconception, by true contents.[12]

Thus, the belief an individual expresses using sentences containing the term "clam brisket," e.g., "I had clam brisket for dinner last night," would, following to this method, be interpreted as a belief about *clisket*, where "clisket" is introduced as having an extension which not only includes that of "brisket," but also covers supposed instances of clam brisket.[13] So, as we would imagine, even though we would use the following sorts of sentences to attribute beliefs to the individual, e.g., "so-and-so believes that he had clam brisket for dinner last night," we would not interpret the word "brisket" literally. Rather, we would interpret the ascription as saying that the individual thought that he had clisket for dinner last night. Burge refers to this as the "object-level" method of reinterpretation.

The other, closely related, method of reinterpreting the thought experiments, the "metalinguistic" method, proceeds from the claim that the individual's incomplete understanding is more accurately described as a metalinguistic error. This method attempts

---

[12] *Ibid.*, p. 549.

[13] This parallels the introduction of the term "tharthritis" which gets introduced in the course of Burge's discussion of the example involving the arthritic patient. "Tharthritis" is introduced as a term whose extension includes that of arthritis and also that of supposed instances in which arthritis is said to occur in the thigh.

to account for the individual's misuse of the particular term. So, for example, the belief an individual expresses by saying "I had clam brisket for dinner last night" involves a simple metalinguistic error--that of thinking that "brisket" applies to certain parts of clams as it does to certain cuts of beef. According to this method, the sentence used to ascribe the belief should be reinterpreted, since it is clear that the individual has simply picked the wrong sentence to express his belief.

As Burge notes, these methods can be applied simultaneously. The meta-linguistic method can be applied in order to account for the sense in which the individual has misconstrued the particular term. On this method, it is supposed that the individual has a false belief about the meaning of the relevant term. So, when the individual uses the sentence "I had clam brisket for dinner last night," it is supposed that the individual has a false belief about the meaning of the word "brisket." And, hence, it is argued, the ascription needs to be reinterpreted in terms of what the subject thought the word meant.

Burge's criticism of these methods of reinterpretation is based on two general claims. First, Burge says, the methods fail to account for the practice of ordinary mentalistic attributions (what we typically say and do when we catch others using words incorrectly). And, second, the reinterpretations urged by the methods are not supported by what the individual would say and do when he realizes that he had been using the particular term incorrectly. I will consider each of these general claims in turn.

So, to begin with the first of these claims, Burge's view is that we do not typically (in ordinary practice) search for true object-level contents, nor do we ordinarily suppose that all of the individual's attitudes involving the misconstrued term involve reference to expressions

at the metalinguistic level. Burge makes three observations, based on examples involving ordinary practice, which are supposed to sway our intuitions against these methods of reinterpretation.

Burge's first observation is that searching for true object-level contents would be non-trivial and *ad hoc*. So, to take the clisket example, it is just not clear, Burge would say, what is to count as clisket. Suppose an individual uses the term "beef brisket" correctly on a number of occasions (or, perhaps, on every occasion on which he uses it). Why should we use "clisket" (or, perhaps, "beef clisket") to characterize his beliefs in cases where he correctly uses "beef brisket"? Burge contends that it is simpler and equally informative to suppose that the individual believed, falsely, that what he had for dinner last night was clam brisket. We do not, on Burge's view, need to reinterpret all of the sentences which contain the term "beef brisket," we just have to keep in mind that some of the patterns of inferences involving the notion of brisket may turn out, upon examination, to display slight deviations from those expected of someone with a complete understanding of the notion.

The second observation is intended to show that concerns about the alleged source of the subject's error do not typically issue in our judgments on the contents of his thoughts. First, Burge notes that the metalinguistic method of reinterpretation does indeed offer a plausible account of how the individual may have come to have the deviant belief. It is not implausible to infer from the individual's utterance of "I had clam brisket for dinner last night," that the individual thinks "brisket" extends to clams. Although the relevant metalinguistic ascriptions may be highly plausible, Burge contends that they do not affect the object-level ascriptions decisively. The individual may indeed believe that "brisket" applies

to certain parts of clams (and does, therefore, have a false metalinguistic belief about "brisket"), but this, by itself, does nothing toward establishing the claim that the individual's belief is not about (*de dicto*) brisket.

The third observation is this. Although the individual has an incomplete understanding of the notion of brisket, he may, as we have already noted, use the word "brisket" correctly on a number of occasions. The individual may have many *prima facie* correct beliefs about beef brisket. He may believe, correctly, that a certain restaurant serves the best beef brisket in town, he may also believe that what his mother cooked every Sunday night was beef brisket and that beef brisket is typically more expensive than pork chops.

We would ordinarily attribute to this individual many unproblematically true (object-level) beliefs about brisket, beliefs that, we would suppose, may be shared by other members of his community. But, if the metalinguistic method were correct, Burge argues, it would be incorrect to claim that the individual shares any (object-level) beliefs with members of his community who have a more complete understanding of the word "brisket."

For example, take a case in which a lawyer's client believes, falsely, that in order to have a contract it is necessary to have a written document. Clearly, the client has a false metalinguistic belief about the meaning of the word "contract." Although the client has this false metalinguistic belief, we would suppose that the client could share many true and unproblematic object-level beliefs with his lawyer (who, presumably, has a more complete understanding of the word "contract"). They could agree, we would imagine, that it was not wise of the client to have signed a mortgage contract without reading all of the fine print.

This is exactly the sort of incomplete understanding which gives rise to Burge's thought experiments. The client obviously has an incomplete understanding of the word "contract" and the lawyer, as we have supposed, has a more complete understanding of the term. Burge's point is that we would use the same ascription to attribute attitudes to both the client and the lawyer, say, "he believes that it was not wise of the client to have signed the mortgage contract without reading all of the small print." In this case, ordinary practice would seem to support the claim that they *both believe that* it was not wise of the client to have signed the contract. The metalinguistic method, Burge notes, would be committed to the counterintuitive claim that the client and the lawyer share only beliefs on these occasions with metalinguistic contents.

Burge's second general claim is that the metalinguistic and the object-level methods of reinterpretation are committed to a highly implausible account of how the individual would react when he discovers that he had been using a term incorrectly. When, for example, the subject learns what brisket is, he does not, Burge contends, typically respond by saying that his views have been misunderstood. Rather, the individual is typically willing to revise his use of the term on the authority of an expert or a reliable source. Moreover, the individual typically admits that the belief he had expressed by saying "I had clam brisket for dinner last night" was false. This suggests that the individual intended to have his words taken literally.

Burge takes this to support the claim that the individual's belief really was about brisket. For, contrary to the object-level method of reinterpretation, the individual does not say or think that his belief had been about clisket. In contrast to the metalinguistic method, the individual does not ordinarily suppose that his belief had been, in any intuitive sense,

concerned with the use of linguistic expressions. The individual is not typically disposed to accept formulations of ascriptions at the metalinguistic level as true accounts of what his belief had been about.

For example, he typically doesn't think that the following sorts of ascriptions capture his belief: "So-and-so believes that the word 'brisket' applies to certain cuts of beef" or "So-and-so believes that the word 'brisket' applies to certain parts of clams." The individual may have these other beliefs. In fact, it is reasonable to suppose that he does have a number of beliefs about the meaning of "brisket." But, on Burge's view, this does not undermine the relevant object-level attributions. All other things being equal, Burge contends, people tend to think at the object level, not the metalinguistic level.

Burge also points to the fact that there are grammatical and semantical differences between formulations of ascriptions at the object level and formulations of ascriptions at the metalinguistic level. He gives the following sort of analysis:

(1) certain parts of clams are covered by the term "brisket"

and

(2) certain parts of clams are brisket

Burge notes that when these clauses are embedded in belief sentences, they express distinct propositions.

Ordinary linguistic practice, Burge says, supports ascriptions at the object level, such as:

(3) So-and-so believes that certain parts of clams are brisket.

In order for the metalinguistic method of reinterpretation to override the ordinary object-level interpretation it is necessary, Burge thinks, to show that there is no reading of (3) that embeds (2). But, Burge argues, evidence upon which ordinary mentalistic attributions are commonly based suggests that correct belief attributions of (3) embed (2) rather than (1). According to Burge, there is simply no evidence which suggests that (3) ever embeds (1) and not (2).

Although in certain contexts the available evidence may equally support embedding (2) and (1)--as Burge notes it is very often the case that there is a relevant metalinguistic belief for every object-level belief--such cases do not undermine the object-level ascription. As Burge says:

> . . . assuming that the object-level and metalinguistic contents are equally attributable, it remains informally plausible that the person's attitudes are different between actual and counterfactual steps in the thought experiment.[14]

The metalinguistic method fails, Burge thinks, in that the metalinguistic and the object-level contents are clearly not equally attributable in describing the contents of the actual and counterfactual individuals, since they have different truth values.

Burge considers three arguments (or argument sketches) which attempt to motivate a non-literal reading (of the sentences used to attribute attitudes) on *a priori* grounds. These arguments are supposed to show that in cases of incomplete understanding, a literal

---

[14] *Ibid.*, p. 552.

interpretation of the sentence the subject uses to express his belief couldn't possibly capture the subject's thought contents.

The first of these arguments is closely related to the metalinguistic method of reinterpretation. Burge sketches the argument as follows.

> One [of the arguments] holds that the content clauses we ascribed must be reinterpreted so as to make reference to words because they clearly concern linguistic matters--or are about language.[15]

This argument attempts to establish the claim that in all of the cases (of incomplete understanding) which give rise to the thought experiments, the beliefs intuitively concern language or are about the use of linguistic expressions. And, as such, the beliefs should be interpreted metalinguistically. For example, it is argued that the belief that someone might express by saying "large overstuffed armchairs are sofas" is intuitively linguistic in the sense that it is a matter of linguistic usage whether large overstuffed armchairs are sofas, and that this is a matter which can be determined simply on the basis of the meanings of the terms which are used. Hence, the argument runs, the belief that the subject expresses by saying "large overstuffed armchairs are sofas" should be interpreted as a metalinguistic belief in virtue of the fact that it is intuitively about language.

Burge has two responses to this argument. First, he claims that not all of the cases which give rise to the thought experiments are explicitly linguistic in nature and, second, even those beliefs which do have linguistic implications, for example, believing that no bachelor is married or believing that contracts are written documents, are not purely

---

[15] *Ibid.*, pp. 552-553.

metalinguistic in form. The beliefs that are expressed by saying "Kirkpatrick is playing a clavichord" and "Milton had severe arthritis in his hands" are not, Burge argues, in any sense linguistic. Moreover, even beliefs which do have an intuitively linguistic nature, such as the belief that no contract is a verbal agreement, are not explicitly metalinguistic. For, as we have already seen, many object-level beliefs have corresponding metalinguistic beliefs, but this alone is not sufficient to undermine the object-level ascriptions.

The second argument that Burge considers may be called the "Argument from Charity." This argument begins with a commonly accepted principle of interpretation, the principle that we should not attribute to rational people irrational or contradictory beliefs. Burge finds this argument unpersuasive, claiming that the sorts of ascriptions which give rise to the thought experiments are perfectly understandable, often the result of "linguistic misinformation." There is, on Burge's view, nothing irrational about believing that one had clam brisket for dinner the other night, or believing that one has arthritis in one's thigh. Furthermore, the Argument from Charity begs the question in that an attitude is uncharitably attributed only if, Burge says, it is first assumed that correct attributions entails complete understanding.

The third argument which Burge examines is called the "Argument from Deviant Speaker Meaning." One might suppose, the argument runs, that when a speaker uses a term incorrectly, he couldn't possibly mean the same thing by his use of the term as we (non-deviant speakers) mean when we use the term correctly. And, consequently, the proposition expressed by the sentence that the deviant speaker uses couldn't possibly capture what he has *in mind*. Burge describes the objection as follows:

. . . we should not attribute contents involving incompletely understood notions because *the individual must mean something different by the misunderstood word than what we non-deviant speakers mean by it.*[16]

There are, Burge thinks, at least two ways to interpret the expression "the individual meant something different by his words." On the one hand, one could think, as Burge himself does, that the expression simply means that the individual only incompletely understood the words he used. On the other hand, one could think that the individual actually had something different in mind. The latter interpretation is where the real force of this objection lies.

Burge thinks that in order to turn this objection into a valid argument it is not only necessary to show that the speaker had something different in mind, it is also necessary to show that the speaker could not have had the attitudes with the commonly attributed contents. That is to say, Burge thinks that in order for this argument to be successful, it is not only necessary to establish the claim that the individual had something different in mind, it is also necessary to show that the literal (or conventional) meaning of the term could not possibly play a role in the subject's thinking as well. Burge is unconvinced by this argument for the following reason:

It does not follow from the assumption that the subject thought that a word means something that it does not (or misapplies the word, or is disposed to misexplain its meaning) that the word cannot be used in literally describing his mental contents. It does not follow from the assumption that a person has in mind something that a word does not denote or express that the word

---

[16] *Ibid.*, p. 553.

cannot occur obliquely (and be interpreted literally) in that-clauses that provide some of his mental contents.[17]

The most this argument could establish, Burge thinks, is that the notion associated with the deviant word may play a role in the speaker's thinking. This claim, Burge thinks, does not undermine the thought experiments. The argument has not shown that the conventional meaning of the term does not play a role in the speaker's thinking as well.

IV: An Objection to Social Externalism

As noted above, Burge's argument for social externalism hinges on the following principles.

(P1) In cases in which a speaker only incompletely understands some concept and makes use of a term which expresses that concept, it is correct to interpret the sentence that the speaker utters (containing that term) in a straightforward and literal way, provided that his misuse of the term is not too deviant, relative to the conventions of his linguistic community.

(P2) In such cases, the content of the individual's belief is given by the proposition expressed by the sentence that the individual uses to express his belief, relative to the linguistic conventions of his community.

In this section, I aim to show that these principles lead to a contradiction. Consequently, we must reject social externalism.

Consider an ordinary, everyday, real-world counterpart to Burge's Twin Earth thought experiment. Imagine a real-world situation in which a single individual, Gulliver, is a

---

[17] *Ibid.*, p. 553.

member of two linguistic communities and where these two communities have different conventions concerning the use of the word-form 'arthritis.' We can suppose that Gulliver is the captain of a single-person sailboat and that he (and he alone) sails from one island community to the other; other than Gulliver, we can suppose that there is no further contact between the two communities. Despite the limited contact, the languages spoken by members of the two communities are quite similar; the only difference involves the use of 'arthritis' and otherwise both languages are identical with English. This difference, however, has never come up in any verbal exchange between Gulliver and any island inhabitant. Gulliver has only had discussions in which 'arthritis' was applied to inflammations of joints, but in linguistic community A 'arthritis' applies just to joints, while in community B 'arthritis' applies to a more general class of ailments (in particular, it applies to allied ailments of the thigh). Notice that the truth conditions for sentences containing 'arthritis' will also be different in the two languages--this is a reflection of the difference in linguistic conventions.

Now we can raise an important difficulty for Burge's view. As a member of linguistic community A, the belief Gulliver expresses by saying "I fear that I have developed arthritis in my thigh" reflects an incomplete understanding of the concept arthritis. Nonetheless, P2 entitles us to say that Gulliver's belief is about his thigh. Indeed, Burge would say that Gulliver has a false belief about arthritis. As a member of linguistic community B, however, the belief Gulliver expresses by saying "I fear my arthritis has spread to my thigh" reflects a more complete understanding of 'arthritis' and, presumably, Burge would say that Gulliver's belief is true. On the face of it, then, Gulliver has a *single* belief with *two*

semantically distinct contents: one of which makes the belief true, the other of which makes the belief false. It is impossible for a single belief to have two contents, however. Thus, we must deny P1 or P2. Since Burge's argument hinges on *both* P1 and P2, this would force the advocate of Burge's argument to relinquish any claim to soundness.

There is, however, a line of reasoning open to the proponent of Burge's argument which would render P1 and P2 consistent. Rather than claiming that Gulliver has a single belief with two distinct contents, one might argue that he has *two distinct beliefs*--one involving the concept of arthritis$_a$ and another involving the concept of arthritis$_b$. At first blush, this seems to be a reasonable suggestion. After all, Gulliver is a member of both linguistic communities; is it not reasonable to suppose that he has distinct beliefs relative to each community? It seems intuitively implausible to suppose that Gulliver has two distinct beliefs *simply because* the sentence he utters gets evaluated with respect to different linguistic conventions. But we need not rest our objection on this.

If, indeed, Gulliver has two distinct beliefs, then it must be the case that either: (a) he has both beliefs all along or (b) at some point in his voyage, he loses one belief and acquires the other. Let us consider option (b) first. Here the claim is that Gulliver's beliefs undergo some sort of change as he travels from one community to the other. Now, how are we supposed to understand this suggestion? Well, P2 would seem to sanction the claim that while Gulliver is in community A, his thoughts are about arthritis$_a$ and that while Gulliver is in community B, his thoughts are about arthritis$_b$. It would be wildly implausible, however, to suppose that as Gulliver approaches one of the communities, the thought appropriate to that community simply pops into existence, while his former thought pops out

of existence. When does it happen? Common sense tells us that our beliefs and desires are relatively stable and persistent over time, not that they pop in and out of existence.

On the contrary, it is natural to suppose that Gulliver would have whatever beliefs he has even on those occasions when he is in neither community--when, for example, he is out to sea. To strengthen our intuitions here, let us suppose that on one unfortunate afternoon, while out to sea, Gulliver's ship is seized by a typhoon; blown off course, he finds himself marooned on an uninhabited island. Now, if, as we have supposed, the contents of Gulliver's thoughts depend on the linguistic conventions of the community in which he happens to find himself, then as long as he is stranded, in isolation from members of either community, it looks as though we are forced to say that the contents of his thoughts are indeterminate. Yet, any theory which commits us to saying that a person's thought contents would be rendered indeterminate if there were no other language users in his or her immediate vicinity must obviously be wrong.

We should be rather hesitant to suppose that Gulliver's thoughts appear (and disappear) magically, as he travels from one community to the other. We should be hesitant because we know that if there is a change in Gulliver's belief system, then there must be something which is causally responsible for that change. So, what needs to be explained here, if one opts for this line, is how it could come to pass that Gulliver loses one belief and acquires the other. For the anti-individualist who is not disposed to accept magic, this account must explain how linguistic conventions could be causally responsible for the supposed change in Gulliver's beliefs, since it is in virtue of the difference in linguistic conventions that Gulliver has been said to have two different beliefs.

*

Let us turn, now, to option (a) and entertain the supposition that Gulliver has both beliefs all along. Notice that on this supposition we are not forced to say that Gulliver's beliefs pop in and out of existence. Indeed, this supposition seems perfectly consistent with our best scientific theories. After all, Gulliver is a member of both linguistic communities, so it is perfectly reasonable to suppose he has developed whatever causal relations are sufficient to entertain arthritis$_a$ thoughts, as well as those which are sufficient to entertain arthritis$_b$ thoughts. But, surely, this cannot be right.

First of all, this will force us to say a number of bizarre things about Gulliver. We must first ask whether in a given community when he utters a sentence using "arthritis," whether he is announcing both of his beliefs or only one of them. If he is announcing both, he certainly does not seem to be aware of it, so we engender immediately a rather bizarre ignorance of thought content. But this doesn't disappear if we suppose he announces one of the two beliefs in one community and the other in the other community, for then he doesn't think he's announcing different beliefs.

Second, notice that there is a single physical state with which both of these beliefs are associated. Generally, we suppose that a belief is a physical state, and we cannot have more than one belief per physical state. A belief is a disposition. Yet it appears that Burge would be committed on this line we're now considering to a single disposition, hence to a single belief which has two contents. That, however, as we have noted, is a contradiction. Moreover, the potential problem is not limited to two contents for a single belief disposition. For we can imagine Gulliver traveling through an archipelago in which each successive

community uses a word slightly differently, although Gulliver imagines the use is the same in each.

It might be objected that we are not committed to saying that the beliefs are identical with a single physical state, for it might be maintained that the belief state is actually a state of the individual plus his environment. But there are two problems with this. First, most externalists recognize that just because an object or state is relationally individuated, it does not follow that the state itself is spread out over the world between the object so individuated and the things in relation to which it is individuated. For example, being a planet is a relational property, but we don't suppose that if something is a planet it is a thing spread out over the whole of the solar system of which it is a part. Likewise, we don't think husbands and wives expand and contract as their distance from one another changes. In addition, we can note that on this view, if one's linguistic community ceased to exist, those beliefs which were dependent on the practices in that community would apparently have to cease to exist as well. But, surely, no one wishes to say that, and certainly not Burge.

It might also be objected that we do not have to individuate beliefs by physical dispositions. That is, it might be said that just as a certain painting might be both my favorite painting and your least favorite painting, so a disposition could be both of two different beliefs. This is a consistent position, and it is the one I think Burge would take. But it would require rejecting what is a fairly well established thesis in the philosophy of mind. And it gives rise to some real puzzles about explaining what we assert. Why did Gulliver utter the words "I've got arthritis"? Standardly, we would say that he believed that he had arthritis and thought that the sentences "I've got arthritis" expressed that belief, and he wanted to express

that belief. But now what do we say? That he uttered those words also because he believed that he had tharthritis and believed that "I've got arthritis" expressed that belief, and he wanted to express that belief? Was there some sort of strange overdetermination? That would be hard to countenance. When Gulliver first travelled to community B, what miraculous change came over him to conjure into existences this alternative state which now plays the same explanatory role as the old one coexisting with it did? Rejecting the view that there is one belief per disposition simply runs us into additional difficulties which we cannot find our way out of.

In addition, let me call your attention to the fact that P2 contains a definite description--P2 says that ascriptions containing embedded content clauses provide *the* content of a speaker's attitude. The use of the definite description plays a crucial role in Burge's argument: it secures the claim that the contents of Smith's attitudes and the contents of Twin Smith's attitude each involve one and only one concept. Now, if it is supposed that Gulliver has attitudes with contents involving both concepts, the definite description in P2 would fail to secure a unique referent. In this case, P2 would be rendered false (or meaningless, depending on one's views on definite descriptions).

We have, thus far, considered three responses to the case of the traveling bilingual: (1) that Gulliver has a single belief with two distinct contents; (2) that Gulliver's thought contents mysteriously change, popping in and out of existence as he travels from one community to the other; and (3) that Gulliver has two distinct beliefs all the while. To this list, let me add a fourth response; one which, I think, is the obvious and intuitively correct response--the individualist explanation. Gulliver has only a single belief; he has simply

picked the wrong words to express his belief in one of the communities, a mistake which is explained by his ignorance of the standard usage of 'arthritis' in one of the languages. The sentence-form "I fear that I have arthritis in my thigh" is true as tokened by Gulliver in community B and false as tokened by Gulliver in community A. This, however, follows trivially from the fact that sentences containing 'arthritis' have different truth conditions in the languages of the two communities.

V: Social Externalism and Skepticism

Let us turn, now, to the question of whether social externalism provides an adequate response to the skeptical challenge. I think it is obvious that it does not. This, however, is not a direct criticism of Burge's theory. At the time Burge wrote "Individualism and the Mental" he did not take himself to be "responding" to the skeptic. Nonetheless, I think it shall be instructive to see exactly where social externalism, billed as a response to skepticism, fails. To this end, let us suppose that social externalism is true.

In assessing the consequences of accepting social externalism for the traditional problem of skepticism about the external world, we should remind ourselves of some important qualifications to the position. First of all, Burge explicitly notes that this conclusion applies to an individual only if the individual intends to be using his worlds in the same way as other members of his linguistic community. Thus we get of course no general claim to the effect that thought content as such depends on the existence of a social community in which an individual is embedded.

What about an individual who intends to be using his words in conformity to a linguistic community when there is no such community? What should we say about a Robinson Crusoe whose former linguistic community has ceased to exist (perhaps all having been the victim of the plague)? Would he cease to be a thinking being all together? Well, obviously not. So the existence of a linguistic community even when an individual intends to use his words in conformity with one is not necessary for him to have thoughts.

But is it necessary for such a community to have ever existed? Suppose our Robinson Crusoe in fact is the only human being that has ever existed, a quantum fluke on a desert island, but that he fancies himself to be a member of a distant linguistic community in a fair land called "England." Is this description incoherent? No, and nothing in Burge's thought experiments suggests that it is. In this case, the case in which an individual thinks falsely that there is some linguistic community to which he is related, which he would describe as community F, and intends to use his words in conformity with what he takes to be the linguistic conventions of F, we would simply say that the thoughts he expresses are fixed by his own dispositions, as if he were the whole of the community, as in fact he is.

In this case, we can see clearly that at best Burge gets a conditional conclusion about the dependence of thought on social relations: if we do not live in a solipsistic world (a world in which there is no linguistic community of which we are members or to which we can or do intelligibly defer in our usage), then, if we intend to use our words in conformity to a linguistic community, and we do not misuse them too egregiously, then our thoughts depend on the social facts about the use of words in the relevant community. Nonetheless, it is clear that this is no use against the skeptic, since to arrive at the conclusion that there is a social

world (and, hence, an external world that we live in), we would have to know that the antecedent clause of this whole conditional obtains. But this expresses a claim about the external world. To assert that we know it is straightforwardly to beg the question against the skeptic.

IV: Individualism and the "Cartesian" Thought Experiments

Although in his early writings Burge does not take himself to be responding to the skeptic, in some of his later work he takes up the skeptical challenge directly. Of particular importance is the anti-skeptical argument he advances in "Cartesian Error and the Objectivity of Perception." His argument proceeds in two stages: in the first stage, he argues that the skeptic is committed to (what he calls) the "individualist assumption"; in the second stage, he argues (a) that the sorts of thought experiments on which traditional skeptical arguments rest fail to provide adequate motivation for this assumption and (b) this assumption is dubious. Much of our critical attention shall focus on whether the skeptic is really committed to the individualist assumption.

What, then, is the individualist assumption? Burge intends the term "individualism" to apply, somewhat roughly, he says, to:

> . . . philosophical treatments that seek to see a person's intentional mental phenomena ultimately and purely in terms of what happens to the person, what occurs within him, and how he responds to his physical environment, without any essential reference to the social context in which he or the interpreter of his mental phenomena are situated.[18]

---

[18] *Ibid.*, p. 555.

Individualistic presuppositions, Burge thinks, have in one way or another pervaded both traditional and modern treatments of the mind. Such presuppositions, Burge contends, are derived from an outdated model which "likens the relation between a person and the contents of his thought to seeing, where seeing is taken to be a kind of direct, immediate experience."[19] On this model, the contents of our thoughts are taken as given, as immediately accessible to us through introspection. And, to state the requisite claim, what is immediately given to the mind is infallible.

Of course, Burge doesn't think that anyone today would seriously entertain such a view--at least, Burge says, in such an "unqualified and robust form." Burge's reasons for discussing this model are not to saddle the individualist with an outdated theory, rather his point is to give some indication of how certain individualist presuppositions have, as he says, infected contemporary theorizing.[20] Although, for example, the Cartesian model of the mind

---

[19] *Ibid.*, p. 555.

[20] Burge has much more to say than can be discussed here about what areas in contemporary philosophy have been infected by individualist presuppositions. A brief comment is, however, in order. Burge thinks that functionalist approaches in the philosophy of mind completely neglect the social aspects that bear on the individuation of thought contents. He says:

> The intentional or semantical role of mental states and events is not a function merely of their functionally specified roles in the individual. The failure of these accounts of intentional mental states and events derives from an underestimation of socially dependent features of cognitive phenomena. (p. 558)

Burge also thinks that his arguments speak against the identity theory. Burge takes his thought experiments to show that it is possible for two individuals to be in identical physiological states and yet differ in their intentionally specified (occurrent) mental states or events.

His argument goes as follows. Suppose that *b* is an event described in some

receives little currency in contemporary philosophical debate, Burge contends that certain aspects of this model have been uncritically accepted, and have even been unreflectively incorporated into contemporary perspectives on the mind.

Burge contends that Cartesian skepticism is the illegitimate heir of this outdated model of the mind. Individualism is a mistake which has pervaded the post-Cartesian tradition. It is a view which, he thinks, has been held, in one form or another, by philosophers as diverse in perspective as Locke, Berkeley, Leibniz, Hume and Husserl, and, more recently, by many contemporary behaviorists and functionalists. What these philosophers all have in common, Burge says, is that "they all think that the nature and individuation of an individual's mental kinds are 'in principle' independent of the nature and individuation of all aspects of the individual's environment."[21]

He offers three formulations of the doctrine of individualism. The first is intended to capture the form of individualism held by contemporary functionalists.

(1)   Individualism is the view that if one fixes those non-intentional physical and functional states and processes of a person's body whose nature is specifiable without reference to conditions beyond the person's bodily surfaces, one has thereby fixed the

---

appropriate physiological language. Suppose that "b" names the same physiological event occurring in the counterfactual subject. Burge's conclusion is this:

> . . . it is plausible that b need not be affected by counterfactual differences in the communal use of the word "arthritis." Actually, the subject thinks that his ankles are stiff from arthritis, while b occurs. But we can conceive of the subject's lacking a thought event that his ankles are stiff from arthritis, while b occurs. Thus in view of our initial premise, b is not identical with the subject's occurrent thought. (p. 560)

[21] Tyler Burge, "Cartesian Error and the Objectivity of Perception," p. 62.

person's intentional mental states and processes--in the sense that they could not be different intentional states and processes from the ones that they are.[22]

Burge takes the second formulation of the doctrine of individualism to be relevant to the non-materialist tradition.

(2)   Individualism is the view that a person's mental states and processes have intrinsic natures, in the strong sense that the nature and correct individuation of those states and processes (including individuation of their intentional content) is independent of any conditions that obtain outside that person's mind.[23]

A central difficulty with this formulation, Burge remarks, trades on the ambiguity of the term "outside" in the expression "conditions that obtain *outside* that person's mind." The problem, Burge says, is not just that the term "outside" is vague when applied to the mind. It is that we seem to have no recognized procedure for determining what lies "outside" the mind. The root of the difficulty, Burge thinks, is that this formulation is "crude at just the wrong point."

The point that needs to be emphasized, when aiming to characterize the doctrine of individualism, Burge claims, is the conception of the mind as being "self-contained." This is what Burge thinks is common to the various forms of individualism taken up by both modern and contemporary philosophers. With this in mind, Burge presents his third and final characterization of the doctrine of individualism.

---

[22]   Tyler Burge, "Cartesian Error and the Objectivity of Perception," p. 62.

[23]   *Ibid.*, p. 63.

(3)   Individualism is the view that an individual person or animal's mental state and event kinds--including the individual's intentional or representational kinds--can in principle be individuated in complete independence of the natures of empirical objects, properties, or relations (excepting those in the individual's own body, on materialist and functionalist views)--and similarly do not depend essentially on the natures of the minds or activities of other (non-divine) individuals. The mental natures of all an individual's mental states and events are such that there is no necessary or other deep individuative relation between the individual's being in states, or undergoing events, with those natures, and the nature of the individual's physical and social environments.[24]

We will return to Burge's characterization of this doctrine later in this chapter. There, we shall have to consider whether the skeptic is really committed to the doctrine of individualism, in any of the above senses.[25] At present, we will focus on the question of whether the Cartesian thought experiments lend credibility to the doctrine of individualism.

The traditional Cartesian thought experiments are intended to motivate an individualistic conception of the mind--specifically, the conception that "one's mental phenomena are in certain fundamental ways independent of the nature of the empirical and social worlds."[26] Burge begins by setting out what he takes to be the strongest case for individualism, claiming that the case for individualism can be strengthened by leaving out the details about dreamers and demons.[27]

---

[24]   *Ibid.*, pp. 63-64.

[25]   It is worth noting that the characterizations of "individualism" that Burge presents are not equavelent. The first is much stronger than the second two, and the second is stronger than the third. In his paper, Burge seems to consider only the first formulation of the individualist assumption.

[26]   Tyler Burge, "Cartesian Error and the Objectivity of Perception," p. 64.

[27]   He thinks, for example, that the case for individualism based on the dream hypothesis is undermined by the fact that "an interpretation of dreams presupposes thoughts in a wakeful state" (p. 64). Moreover, he notes "the fact that one is being deceived or fooled is of critical importance to any discussion of the relevance of the hypothesis to

> I will construe Descartes as capitalizing on the causal gap that we tend to assume there is between the world and its effects on us: different causes could have produced 'the same' effects, certainly the same physical effects on our sense organs. I will interpret him as conceiving a person as radically mistaken about the nature of the empirical world. I shall see him as imagining that there is something causing the given person's mental goings on, but as imagining that the entities that lie at the ends of relevant causal chains (and perhaps the causal laws) are very different from what the person thinks.[28]

He refers to this as the "Cartesian Hypothesis." A central aim of Burge's paper is to show that the Cartesian Hypothesis does not, contrary to much of what has been assumed in the post-Cartesian tradition, lend any direct support to what he calls the "doctrine of individualism."

This hypothesis, Burge says, rests on two claims: an epistemic claim and a claim about causation. He argues that the claim about causation is not sufficient to establish individualism: "The possibility that very different causal antecedents could issue in the same physical effects on the individual's body, and perhaps even issue in the same phenomenological mental phenomena, is used as a component in my previous arguments against individualism."[29]

Burge's point is that "both the Cartesian thought experiments and my anti-individualistic arguments make use of the possibility that different causal antecedents could have the same effects on the person's surfaces."[30] Burge takes this to show that if

---

individualism" (p. 64).

[28] *Ibid.*, p. 65.

[29] *Ibid.*, p. 65. The article to which Burge is refering is "Individualism and Psychology."

[30] *Ibid.*, p. 66.

individualism does follow from the Cartesian thought experiments, it must be the epistemic claim (and not the causal claim) that is doing the work.

With this said, he turns his attention to the epistemic claim. The epistemic claim is simply the claim that we could be drastically mistaken about the external world. Until this claim is turned into a claim about the difficulty of justifying one's beliefs, Burge notes that it has no skeptical force. Thus, he remarks:

> So let us assume that we know or have reasonable beliefs about what the empirical world is like. And let us grant that the Cartesian thought experiments show that we could be radically mistaken in these beliefs. That is, it is epistemically possible that the world be, or have been, very different from the way we reasonably think it is, or even know it is. This may be seen as a concession that we are deeply fallible. We can imagine being, and perhaps even being shown to be, pretty spectacularly wrong. But it is not a concession that there is reason to think that the beliefs that we are conceding might 'in principle' be wrong really are wrong, or even unjustified.[31]

Burge's next move is to show that the inference from the epistemic claim to the doctrine of individualism is one which begs the question at hand.

> So what follows from this concession of the Cartesian epistemic possibility? Nothing immediate that favors individualism. We took our thoughts about the world as a *given* and conceded that they might be radically mistaken. But we conceded nothing about how our thoughts about the world are determined to be what they are. That is the issue before us. To assume that the epistemological intuitions occasioned by the Cartesian thought experiments support individualism is to make a step that needs justification. It is to beg precisely the question at issue.[32]

---

[31] *Ibid.*, p. 66.

[32] *Ibid.*, p. 66.

What stands in need of justification here, Burge thinks, is the move from "Things might have been radically otherwise without our surfaces being differently affected; and relative to these imagined circumstances, our (actual) thoughts would be subject to numerous and radical errors" to the claim that "Things might have been radically otherwise and our thoughts and minds *would* remain just as they are."[33]

I want to pause here to evaluate both of these claims: the causal claim and the epistemic claim. Concerning the first, I will simply note that there is little textual evidence which supports Burge's contention that Descartes was advancing a claim about causation. Moreover, none of the assumptions of philosophical skepticism involve claims about causation.

Let us move onto the second claim. How exactly are we supposed to understand this claim? In the quotation above Burge says "let us assume that we know or have reasonable beliefs about what the empirical world is like." Let us say that we know that p, where "p" expresses a set of propositions about the external world. In other words, let us assume that skepticism about the external world is false. Yet, in the next sentence he says "And let us grant that the Cartesian thought experiments show that we could be radically mistaken in these beliefs." Clarifying this, Burge says, "That is, it is epistemically possible that the world be, or have been, very different from the way we reasonably think it is, or even know it is."

What is puzzling here is Burge's use of the term "epistemically possible." Something is epistemically possible just in case it is not incompatible with what we know or justifiably

---

[33] *Ibid.*, p. 67; italics added.

believe. If we know that skepticism is false, then we are in an epistemic position know that all possibilities that are consistent with the truth of skepticism are false. So, if we know that skepticism is false, it is not epistemically possible that the world is radically different from what we take it to be. So far as I can see, Burge's claim here is incoherent.

Now, as I mentioned in the introductory section of this chapter, Burge is not concerned with the problem of skepticism about the external world. He is concerned to show that the Cartesian thought experiments do not lend any direct support to individualism. Our interest in Burge's discussion lies in the fact that the skeptic must assume something very similar to individualism in order to get his argument up and running--i.e., the skeptic must assume that our minds are logically independent of the external world.

I say "something very similar" because it should be clear that the skeptic does not need individualism in the strong sense in which Burge outlines it: the individualist may need the claim that "our thoughts and minds *would* remain just as they are," but the skeptic only needs the claim that our thoughts and minds *could* remain just as they are. Burge has failed to establish the claim that the skeptic is committed to the individualist assumption. Thus, we can note that the first stage of Burge's argument is unsuccessful.

Nonetheless, Burge takes what he has said thus far "to undermine the sense that the Cartesian thought experiments provide simple, direct support for individualism."[34] As Burge sees it, the mistake rests on one or both of the following conflations: the first is "conflating questions of counterfactually evaluating one's thoughts with questions of what thoughts one would think if one were in the counterfactual situation" and the second is "conflating the fact

---

[34] *Ibid.*, p. 68.

that we are authoritative about our actual thoughts, and would be authoritative about what our thoughts would be in any (relevant) counterfactual situation, with the claim that we are actually authoritative about certain thoughts that we would be thinking regardless of what actual or counterfactual situation we would be in."[35]

If what I have said above is correct, it follows that even if Burge is right that such conflations are by those who argue for what Burge calls "individualism" (what I shall call "strong individualism") this does not show that the skeptic's assumption is incorrect, since the skeptic relies on a weaker claim.[36] This skeptic just needs to argue that our thoughts *could* be the same even though the world be very different from what we take it to be, not that we *would* have the same thoughts even though the world be very different from what we take it to be.

But one might also think that Burge's charge that some conflation is involved in the argument for strong individualism also applies to thought experiments designed to support weak individualism, the claim that the skeptic actually needs. After all, the skeptic does need some reason to think that our thoughts could be different. On what grounds does the skeptic advance this supposition? Presumably, it is that we can coherently conceive, for example, that beings like us with thoughts like ours (with respect to thoughts attributed using general terms and singular terms that refer only to ourselves)[37] could exist in a world very different from the way we believe the world to be.

---

[35] *Ibid.*, p. 68.

[36] Indeed, the claim that the skeptic needs is much weaker than any of the three formulations of individualism that Burge characterizes.

[37] For discussion of this point, see Ludwig (1996a) and (1996b).

Might it not be charged that the appearance of conceivability here is due to our taking our actual thoughts as given and imagining a world that is very different from the world our thoughts would be of if veridical? It might indeed, but is that at all the way I have described the thing we are trying to imagine? Here is what I imagine: there is a being just like me with respect to those thoughts which I have access to without recourse to the nature of my environment. That being is not me, but exists in a world in which, sadly, her thoughts about her world are mostly false.

To make this vivid, we could imagine an observer in that world who has true beliefs (and who knows what her beliefs are) seeing how mistaken she is, though this is not necessary. Is this conceivable? I believe it is. But can I be accused of the conflation Burge thinks has deluded philosophers here? Surely not, for I have been careful to state exactly what it is that I am imagining, and it is not a world in which my actual thoughts are false, but a world in which someone who has thoughts like us might be mostly mistaken.

We may safely conclude, then, that Burge's reflections on what he calls the "Cartesian thought experiments" do not, in fact, succeed in showing that there is any irredeemably faulty reasoning involved.

V: Burge's Argument from the Objectivity of Perception

Thus far, Burge has argued for a limited conclusion--that the Cartesian thought experiments fail to establish the doctrine of individualism. Let us now take up Burge's argument for the claim that individualism is false. His argument is based on considerations about perception.

His argument rests on three premises. The first premise states that "our perceptual experience represents or is about objects, properties, and relations that are *objective*."[38] The second premise holds that "we have perceptual representations (or perceptual states with contents) that *specify* particular objective types of objects, properties, or relations *as such*."[39] The third premise states that "some perceptual types that specify objective types of objects, properties, and relations as such do so partly because of relations that hold between the perceiver (or at least members of the perceiver's species) and instances of those objective types."[40] Given these three premises, it can be established, Burge claims, that "individualism is not true for perceptual representations."[41]

Burge asks us to consider the following thought experiment. Imagine an individual, S, who has perceptual experience with certain objective entities. Under normal conditions, S perceives tokens of these objective entities (call them "O's") correctly (as O's). In certain aberrant cases, S misperceives instances of O's as instances of C's. (We can suppose that C's are superficially similar to O's.) Burge invites us to imagine that S's discriminative abilities with respect to C's and O's are such that at the time of the misrepresentation S cannot discriminate instances of C's from instances of O's.

Now, holding fixed S's physical states and discriminative abilities (described non-intentionally and independently of the nature of S's environment), we are asked to imagine a counterfactual situation in which there are no O's. Further, we are asked to imagine that

---

[38] *Ibid.*, p. 69.

[39] *Ibid.*, p. 69.

[40] *Ibid.*, p. 70.

[41] *Ibid.*, p. 73.

S's surface irritations which are actually caused by normal perceptions of O's are counterfactually caused by normal perceptions of C's. We may also suppose, Burge says, that members of S's species have adapted to their environment and, perhaps, that the laws of optics are different in the counterfactual situation. Note this point, as I will come back to it later. This assumption--that the beings in question are evolved beings, suited to their environements--I believe helps give plausibility to the intuitions Burge hopes to elicit--but at a cost.

Burge takes it to follow from premises two and three that in the actual situation the contents of S's perceptual experience are about O's and in the counterfactual situation the contents of S's perceptual experience are about C's. Thus, Burge concludes "since the objective entities that the person normally interacts with--and perceives as such--differ between actual and counterfactual situations, and since the laws explaining these interactions also differ, the perceptual intentional types of the person also differ."[42] In other words, "the person's physical states, discriminative abilities, and perhaps purely phenomenological (non-intentional) states remain the same between the two situations. So, the person's intentional perceptual types are not individualistically individuated."[43]

To clarify the argument, Burge offers the following example. Suppose that the entities in question are relatively small and do not affect S's ability to adapt to her environment. Let the O's be shadows of a certain size on a contoured surface and let the C's be similarly sized cracks. Such shadows are common in S's environment and S interacts

---

[42] *Ibid.*, p. 74.

[43] *Ibid.*, pp. 74-75.

with them on a regular basis, but the corresponding cracks are fairly uncommon and S rarely interacts with them.

Now suppose that at time t S is unable to distinguish the commonly seen shadows from the cracks she encounters on occasion, though S is quite able to distinguish such shadows from other relevantly similar objective objects in her environment. Now, suppose that at time t S encounters one of the cracks in question. It is clear, Burge claims, that we would describe this as a case in which S misperceives the crack as a shadow.

Next, Burge asks us to consider the following counterfactual situation. We hold constant S's physical states and discriminative abilities (non-intentionally described), and imagine that S's environment is just as it was in the actual situation, with the exception that where there were shadows of the relevant size and shape, there are relevantly similar cracks. In the counterfactual situation, S's perceptual experience is caused by an appropriately sized crack and not a shadow. Burge contends that in this case S is perceiving a crack correctly as a crack. He concludes that the contents of perceptual states are not individualistically individuated.

Before moving on to evaluate Burge's argument, it will be useful to put the argument in explicit form.

[1]   For any individual P, if P inhabits a world in which there are similar objective
       types O and C, and P's perceptual experiences of type E are in a normal
       perceptual conditions typically caused by tokens of O, then it is possible for P
       to misperceive C's as O's. (premise)

[2]   If P had inhabited a world which contains only objects of type C, and was non-
       relationally and non-intentionally the same, P would not be able to misperceive
       C's as O's. (premise)

[3]   Therefore, for any P, if P inhabits a world in which there are similar objective types O and C, and P's perceptual experiences of type E are in a normal perceptual conditions typically caused by tokens of O, then the content of a perceptual experience is given by the type of object, tokens of which normally cause perceptual experience of such entities. (from 1 and 2)

[4]   Therefore, for any P, if P inhabits a world in which there are similar objective types O and C, and P's perceptual experiences of type E are in a normal perceptual conditions typically caused by tokens of O, then the content of a being's perceptual experience is determined in part by facts about that beings physical environment. (from 3)

Burge takes it to follow from [4] that "individualism" is mistaken, since "individualism" entails that [4] is false and he takes himself to have established [4] by the foregoing argument.

VIII: An Objection to Perceptual Externalism

In his "Comments on Burge," Robert Matthews has suggested an alternative explanation of the thought experiment. "Given that the organism does not discriminate cracks from shadows, even though it is supposedly capable of doing so, one could well argue that this organism perceives cracks and shadows as instances of one and the same entity."[44] Indeed, it may very well be that S perceives shadows in the actual situation and cracks in the counterfactual situation, not as shadows and cracks, respectively, but as instances of a more general type of objective entity, call them "cradows." As he notes, unless Burge can provide

---

[44]   *Ibid.*, p. 83.

some principled motivation for ruling out this possibility, his argument is less than conclusive.

IX: Perceptual Externalism and Skepticism

Nonetheless, let us waive the foregoing objections and suppose, for the sake of argument, that perceptual externalism is true. What follows from perceptual externalism? In this section, I argue that nothing of any anti-skeptical force follows from perceptual externalism.

First, we can note that the conclusion of Burge's argument for perceptual externalism takes the form of a conditional--for any P, *if* P inhabits a world in which there are similar objective types O and C, and P's perceptual experiences of type E are in a normal perceptual conditions typically caused by tokens of O, *then* the content of a being's perceptual experience is determined in part by facts about that beings physical environment. To actually get the claim that "individualism" is mistaken, Burge needs to discharge the antecedent of this conditional. Yet do do so would obviously beg the question against the skeptic.

Aside from this difficulty, even if Burge's argument were successful, the most he could establish is a conditional conclusion, whose antecedent involves some claim about the external world. Why is that? First, I want to note that if I am right in what I have said above about the thoughts experiments relying on the assumption that the beings in question are evolved beings, it is clear that the conclusion that is forthcoming is conditional on an empirical hypothesis, that is, the conclusion is, at best: if we are evolved beings, then our

thoughts are about things that regularly cause them. Nonetheless, I do not want to rest my objection on this observation, even though I think it is correct.

Even if we granted, just for the sake of argument, that our intuitions told us that regardless of whether the beings in question were evolved or not, that if their experiences were caused (appropriately) by certain things in their spatial environment, then their experiences would be about those things. Would this be enough to defeat radical skepticism? I think the answer is that it would not. For these would be intuitions about what thoughts beings would have if their world were a spatial one. It would not establish the claim that to have perceptual experiences with representation content at all they had to be in a spatial world.

At best, such intuitions would sanction the claim that if we are in a spatial world, then our experiences are about what regularly cause them. But this is of no use against the skeptic, because to use this is to infer that most of our perceptual experiences are caused by what they are about would require us to establish first that we are in a spatial world. If it is possible that we could have the experiences we do in a non-spatial world, as the skeptic maintains, then to assume that we are in a spatial world begs the question against the skeptic. It assumes some knowledge that falls within the scope of the skeptic's argument.

My point here is about the limitations of the sort of argument that Burge advances, limitations which most philosophers routinely overlook. In considering a thought experiment in which one considers two circumstances S&O and S&C with a common element, S, and noticing that some further element of sort M intuitively is thought to vary with the variance from O to C, we cannot straightaway conclude that the kind of fact illustrated by O and C

(suppose these are both facts of type R) is necessary for the element in question. For we have not controlled for the relevance of S. It may be that M may occur in the absence of S, though if S is in place what determinate M property obtains depends on what R properties obtain. In the present case, S = being in a spatial world, O = having experiences of type E caused by shadows, C = having experiences of type E caused by cracks, and M = having a mental state characterized by representational content. Even if relative to S, M varies with variance in O and C, this doesn't establish that S&X for some X is necessary for instances of M.

Might opponents of the logical independence of thought content from the world try to argue that causal interaction with spatial objects is necessary for thought? Yes. But not by the sort of thought experiment typically employed. Some other argument must be advanced. But is it just intuitively obvious that one cannot have thoughts with certain content unless one is in causal interaction with spatial objects? Not at all, so far as I can tell, and it is significant that this is not how the thought experiments typically proceed. After all, our intuitions do not tell us directly that there could not be a solipsistic thinker whose thoughts evolve according to internal laws and not in virtue of causal interaction with anything else.

I conclude that perceptual externalism, as developed by one of its leading proponents, Tyler Burge, does not provide the basis for a satisfactory response to skepticism, because the arguments advanced for it at best yield a conclusion conditional in form whose antecedent is about certain general conditions obtaining in the external world, so that its deployment against skepticism would have to rest on prior knowledge of facts about the external world-- facts which fall within the scope of the skeptic's negative assessment of our knowledge.

PUTNAM'S ARGUMENTS FROM
THE CAUSAL THEORY OF REFERENCE

I. Introduction

In this chapter we take up a series of arguments advanced by Hilary Putnam which are all based on general considerations about the conditions necessary for meaning/ reference. Our aim in this chapter is to answer two questions. First, is Putnam's account of meaning/reference correct? Second, if this theory is correct, what relevance (if any) does it have for resolving the problem of skepticism?

Here is how we shall proceed. In the next section, we review Putnam's attack on the traditional Fregean conception of meaning and evaluate the account of meaning he offers in "The Meaning of 'Meaning'." In section III, we turn to a discussion of the anti-skeptical argument he presents in the first chapter of *Reason, Truth and History* to the effect that we can know that we are not brains in a vat. In section IV, we shall evaluate Putnam's anti-skeptical argument in light of his more recent work. In particular, in his later work, Putnam gives up the notion of metaphysical necessity. What we shall need to do is evaluate the force of the Causal Theory of Meaning (CTR) in light of this concession. Finally, in section V, I consider a number of arguments which suggest that there are empirical suppositions in Putnam's though experiments. I conclude that Putnam's externalism does not provide an adequate response to the skeptical challenge.

II. A Theory of Meaning

By 1975, the year Putnam published the "Meaning of 'Meaning,'" the traditional Fregean conception of meaning had already come under attack, most conspicuously by Saul Kripke. In *Naming and Necessity*, Kripke argued that the reference of a proper name is not fixed by a single description or a set of descriptions; rather, reference is fixed by a historical causal chain extending from the "baptism" of the name up through its current use. He also argued that the extensions of natural kind terms are not fixed by descriptions, but depend on the actual physical constitutions of the natural kinds to which the terms are applied. The suggestion here, which would later be taken up in Putnam's work, is that the meanings of many of the terms we use depend in important ways on the nature of our environment and our causal interaction with it.

This suggestion constitutes a radical break from the traditional Fregean conception of meaning. On the Fregean conception of meaning, there are two components to the meaning of a term: extension and intension. The extension of a term is the set of things of which the term is true--e.g., 'rabbit' is true of each sundry rabbit, so its extension is the set of rabbits. It soon became clear to the advocates of this theory that there must be more to the meaning of a term than just its extension, since two terms can have the same extension and yet differ in meaning, as Quine's time-worn example involving "creature with a heart" and "creature with a kidney" serves to establish. This additional element came to be known as the "intension" of a term. Proponents of the Fregean conception of meaning met with considerable difficulty in their attempt to provide a clear and informative analysis of the

notion of intension; nonetheless, we can say, roughly, that the intension of a term is the concept associated with it.

In "The Meaning of 'Meaning,'" Putnam presents a critique of the traditional theory of meaning. He argues that this theory rests on two assumptions which had hitherto been uncritically accepted, namely:

(I)   That knowing the meaning of a term is just a matter of being in a certain psychological state (in the sense of 'psychological state,' in which states of memory and psychological dispositions are 'psychological states'; no one thought that knowing the meaning of a word was a continuous state of consciousness . . .).

(II)   That the meaning of a term (in the sense of 'intension') determines its extension (in the sense that sameness of intension entails sameness of extension).[1]

He aims to show that "these two assumptions are not jointly satisfied by any notion, let alone the notion of meaning," and hence that the "traditional concept of meaning is a concept which rests on a false theory."[2]

To see how the argument works, it is necessary to get clear on the notion of "psychological state" at work in assumption (I). This assumption, Putnam claims, is constrained by what he calls "methodological solipsism."

This is the assumption that no psychological state, properly so-called, presupposes the existence of any individual other than the subject to whom that state is ascribed. (In fact, the assumption was that no psychological state presupposes the existence of the subject's body even: if P is a psychological

---

[1] H. Putnam, "The Meaning of `Meaning'," p. 219.

[2] *Ibid.*, p. 219.

state, properly so-called, then it must be logically possible for a 'disembodied mind' to be in P.)[3]

Psychological states which meet this restriction are termed "narrow psychological states" and those that do not are called "wide psychological states." The notion of psychological state at work in assumption (I), Putnam says, is that of a narrow psychological state. So, assumption (I) says that knowing the meanings of the terms we use is simply a matter of being in a certain narrow psychological state: for any subject S, S's knowing the meanings of the terms she uses presupposes only her own existence and not the existence of other speakers, her own body or, we might add, the external world.

It follows from assumptions (I) and (II) that psychological state together with the world determines extension. For if knowing the meaning of a term is being in a certain narrow psychological state, then being in that state determines its meaning (I). But if meaning determines extension (II), then by transitivity of determining it follows that narrow psychological state determines extension. With this conclusion secured, Putnam sets up the target of his attack as follows.

[1]   For any two terms, A and B, if A and B are extensionally distinct, then necessarily knowing the intension of A is a different narrow psychological state than knowing the intension of B. (from I and II)

[2]   For any speaker S, S can be said to grasp the intension of a given term *only if* S is aware that the intension grasped is the intension of that term,[4] and S's

---

[3]*Ibid.*, p. 220.

[4]There is a questionable move here--it would seem that knowing that the intension is the intension of a given term would already require reference to, say, how the majority of others in one's community use terms or what the terms really refer to; for we each

knowing the intension of that term amounts to S's knowing a set of necessary and sufficient conditions for membership in the extension of that term. (premise)

[3]   Let I_1 and I_2 be any two distinct intensions of a term T. The psychological state of knowing that I_1 is the intension of T is different from the psychological state of knowing that I_2 is the intension of T. (from 1)

[4]   Thus, it is logically impossible for there to exist two distinct possible worlds, W_1 and W_2, such that S's narrow psychological state remains constant and yet in W_1 S knows that I_1 is the intension of T and in W_2 S knows that I_2 is the intension of T.[5] (from 1, 2 and 3)

Putnam argues that [4] is false--"it is possible for two speakers to be in exactly the same psychological state (in the narrow sense), even though the extension of the term A in the idiolect of the one is different from the extension of the term A in the idiolect of the other."[6] And since [4] follows from assumptions [I] and [II], (at least) one of these assumptions must be false. In the end, Putnam concludes that assumption (II) is false.

His argument for this conclusion proceeds in two stages. The first stage consists of his famous Twin Earth thought experiment, the purpose of which is to show that narrow psychological state does not, contrary to the traditional Fregean assumptions, determine extension. This leaves three options: (a) that psychological state does not determine intension (that assumption (I) is false), or (b) that intension does not determine extension

make the presumption that the intension we grasp is the "appropriate" ("correct" ?) intension. The point is that we cannot discover our own mistakes by ourselves. Another way to put the point is that if we are really making the assumption of methodological solipsism, then we simply get our intensions by stipulation. As long as we are consistent, we can't go wrong.

[5]It would seem that Putnam intends the expressions "the same intension as" and "the same meaning as" to be equivalent.

[6]*Ibid.*, p. 222.

(that assumption (II) is false), or (c) that one or the other speakers doesn't know the meaning of the term in his idiolect. In the second stage of his argument, Putnam aims to show that option (a) is unacceptable, and hence (rejecting (c)) that we must opt for (b). Combining these two stages, Putnam draws the conclusion that intension does not determine extension or, as he has put the point, that "meanings just ain't in the head."[7] His conclusion is based on his now famous Twin Earth thought experiment.

III. Putnam's Argument from "Twin Earth"

Putnam asks us to suppose that at the far end of the galaxy there exists a planet which is remarkably similar to Earth called Twin Earth. The inhabitants of Twin Earth, we are asked to suppose, speak what sounds like a dialect of English.[8] There are, however, subtle differences between Earth and Twin Earth. The difference on which Putnam's thought experiment turns is that what the Twin Earthlings call *water* has a different chemical structure than the stuff we call water--XYZ and not H_2O. The difference in chemical structure is not detectable by the naked eye; in fact, it would take a number of sophisticated laboratory tests to distinguish between the two substances.

Putnam's claim is that if a spaceship from Earth ever landed on Twin Earth, the crew members would suppose, initially, at least, that "water" had the same extension on Twin Earth as it does on Earth (and that "water" has the same meaning on Twin Earth as it does

---

[7]*Ibid.*, p. 227.

[8]From now on I will in fact call it a dialect of English, for the sake of convenience, though since it has no historical connections with English, this is literally false.

on Earth). They would soon discover their mistake, however. For their mistake would be corrected once it was determined that the stuff surrounding them is composed of XYZ and not H_2O. We can imagine their science officer reporting "On Twin Earth the word 'water' means XYZ."[9] Her report would be correct, Putnam contends: on Earth the word "water" means H_2O and on Twin Earth "water" means XYZ.

Putnam realizes that this thought experiment will not show that it is possible for two speakers to be in exactly the same narrow psychological state even though the extension of the term "water" is different. This is because any difference in extension can be readily attributed to the difference in narrow psychological state--the typical earthling will have different beliefs about the meaning of "water" than the typical twin earthling. Dialectically, Putnam has to set up the thought experiment so that the narrow psychological states of our imagined protagonists remain constant, while the extensions of the terms they use differ. To accomplish this end, Putnam varies the thought experiment slightly.

He asks us to roll back the time until about 1750 before the development of chemistry, on Earth or Twin Earth. In 1750, no one (on Earth or Twin Earth), knew the chemical structure of the substance called "water." Let us imagine two such speakers. Let $Oscar_{TE}$ be a typical Twin Earth speaker and let $Oscar_E$ be his doppelgänger on Earth. In the case we have described, it is clear that neither $Oscar_{TE}$, nor $Oscar_E$ would know the chemical structure of the stuff he calls "water." Yet, as Putnam notes, there has been no change in the extension of the term "water" on Earth or Twin Earth since 1750: its extension on Earth was the set of molecules composed of H_2O and its extension on Twin Earth was the set of

---

[9]*Ibid.*, p. 223.

molecules composed of XYZ.[10] So, although Oscar$_{TE}$ and Oscar$_E$ are in the same narrow

psychological state, their use of the term "water" secures distinct extensions.[11] Thus, Putnam

concludes, "the extension of the term 'water' . . . is *not* a function of the psychological state

of the speaker by itself."[12] Now, if this is true, then, if the speaker does succeed in referring,

then it follows that there must be something in addition to a speaker's narrow psychological

---

[10]A key assumption that Putnam is concerned to defend is that the meaning of the term "water" has not changed since 1750 (on Earth or Twin Earth). To support this assumption, Putnam gives the following analysis of "x is the same liquid as y" (or "x is the same$_L$ as y"). Putnam asks us to consider a case of ostensive definition: I point to a glass and say "This is water." There are, Putnam notes, certain empirical presuppositions which must be satisfied if my ostensive definition is to come off, the most significant of which is that the stuff I point to must bear the same$_L$ relation to the stuff that speakers in my linguistic community, including myself, call water. Should it be discovered that the stuff I am pointing to is, say, a glass of grain alcohol and not a glass of water, then I would not expect my ostensive definition to be accepted. My ostensive definition can come off, Putnam claims, only if the stuff ostended to really is water.

The "same$_L$" relation, Putnam notes, is a theoretical relation in that "whether something is or is not the same liquid as *this* may take an indeterminate amount of scientific investigation to determine" (p. 225). "Moreover," Putnam continues, "even if a 'definite' answer has been obtained either through scientific investigation or through the application of some 'common sense' test, the answer is defeasible: future investigation might reverse even the most 'certain' example" (p. 225).

[11]Putnam states his conclusion as follows:

> Oscar$_{TE}$ and Oscar$_E$ "understood the term 'water' differently in 1750 although they were in the same [narrow] psychological state, and although, given the state of science at the time, it would have taken their scientific communities about fifty years to discover that they understood the term 'water' differently. Thus the extension of the term 'water' is . . . not a function of the psychological state of the speaker by itself" (p. 224).

This is a contentious way to state the conclusion. It doesn't follow that they understood the term 'water' differently. Putnam is only entitled to the claim that sameness of narrow psychological state does not imply sameness of extension. To get the further conclusion (that they understood the term differently), he must assume they understand the terms completely.

[12]*Ibid.*, p. 224.

state that "completes the job of reference." Putnam's contention is that it is the world itself which completes the job of reference.

His explanation is that natural kind terms such as "water" have a hidden indexical element. To see this, consider the case of ostensive definition. One way to teach someone how to use the word "water" is to simply point to a glass of water and say "*this* is water." The act of giving an ostensive definition can come off only if certain empirical conditions obtain. For example, I would not intend for my ostensive definition to be accepted, if the glass I had pointed to happened to contain gin rather than water. More generally, when I point to a glass of liquid, and say "this is water," I am assuming "that the body of liquid I am pointing to bears a certain sameness relation (say, *x is the same liquid as y*, or *x is the same$_L$ as y*) to most of the stuff I and other speakers in my linguistic community have on other occasions called 'water'."[13] This suggests that natural kind terms such as "water" are rigid, to put the point loosely--a given substance (in any possible world) is water only if it bears the same$_L$ relation to the stuff we call water. Since what we call water is H_2O, it follows that water is H_2O in all possible worlds. Putnam's point is that reference to the actual stuff makes a difference in determining what its extension is.

Returning to Putnam's Twin Earth thought experiment, the point is that the stuff Twin Earthlings call "water" isn't water. Indeed, Putnam's contention is that there is no logically possible scenario in which water has a different chemical structure. "Once we have discovered that water (in the actual world) is H_2O, nothing counts as a possible world in

---

[13]*Ibid.*, p. 225.

which water isn't H_2O. In particular, if a 'logically possible' statement is one that holds in some 'logically possible world,' it isn't logically possible that water isn't H_2O."[14]

This is an incredible claim, given that it is conceivable that future discoveries may force us to revise our current scientific theories about the chemical composition of natural kinds. Putnam is well aware of the peculiarity here: "we can perfectly well imagine having experiences that would convince us (and make it rational to believe that) water isn't H_2O. In that sense, it is conceivable that water isn't H_2O. It is conceivable but it isn't logically possible! Conceivability is no proof of logical possibility."[15] Another way to put this point is that it is epistemically possible that water isn't H_2O, but it is not metaphysically possible that water isn't H_2O.

Intuitively, Putnam's claim is that to get to metaphysical possibility, we have to take into account the way the world is; for as Putnam goes on to explain, a given substance is not of substance-kind S simply in virtue of satisfying the current scientific criteria for being S, a given substance is of substance-kind S in virtue of its real physio-chemical structure. However, this is not to suggest that it is metaphysically necessary that all the stuffs we call by a given substance term all have a common physio-chemical structure--e.g., "jade" applies to both jadeite and nephrite.

The point here is that if our environment had contained both H_2O and XYZ, then it would be correct to say that there are two kinds of water.[16] And so it would be correct to say

---

[14]*Ibid.*, p. 233.

[15]*Ibid.*, p. 233.

[16]In fact, in a sense this is borne out, since we count H_2O and O_2O (heavy water) as water.

that the stuff on Twin Earth is water; in particular, it would be correct to say that "it turned

out to be the XYZ kind of water."[17] (If scientists wake up tomorrow morning and discover

some samples of XYZ, say, near the depths of the ocean floor, then perhaps we would have

to revise our supposition that it is metaphysically impossible for water to be XYZ.) This

shows that Putnam is thinking that the way the world is constrains what (we can say) is

metaphysically possible.

Returning to the theory of meaning, Putnam summarizes his criticism of the

traditional Fregean view in the following terms.

> We have now seen that the extension of a term is not fixed by a concept that
> the individual speaker has in his head [that is, a speaker's narrow
> psychological state], and this is true both because extension is, in general,
> determined socially--there is a division of linguistic labor as much as of 'real'
> labor--and because extension is, in part, determined indexically. The
> extension of our terms depends upon the actual nature of the particular things
> that serve as paradigms, and this actual nature is not, in general, fully known
> to the speaker. Traditional semantic theory leaves out only two contributions
> to the determination of extension--the contribution of society and the
> contribution of the real world![18]

This completes the first stage of Putnam's argument--that narrow psychological state does

not determine extension. As noted above, this leaves us with three options: (a) that

psychological state does not determine intension, (b) that intension does not determine

extension and (c) that one or the other speaker does not know the meaning of the term in his

idiolect.

---

[17]*Ibid.*, p. 241.

[18]*Ibid.*, p. 245.

Let us turn now to the second stage of his argument, which is supposed to show that we should opt for (b), rejecting the claim that intension determines extension. His argument is based on the following case. Suppose that I can't tell the difference between an elm and a beech. It is nonetheless clear that when I use the term "elm," I manage to refer to elms, and when I use the word "beech," I manage to refer to beeches. This is so, despite the fact that my concept of an elm is no different from my concept of a beech. This shows that it is not in virtue of the concept I have in mind head (that is, in virtue of my narrow psychological state) that I succeed in referring to elms when I use the word "elm."

Now, suppose that on Twin Earth, elms and beeches are switched--i.e., for every elm (beech) on Earth, there is a beech (elm) in corresponding location on Twin Earth. If this were the case, then we would not say that "elm" has the same meaning on Earth and Twin Earth; rather, we would say that the term "elm" as spoken by my doppelgänger means precisely what I mean when I use the term "beech." This shows that the notion of meaning is far more objective than the traditional account would lead us to believe. To capture this objectivity, Putnam proposes to "identify 'meaning' with an ordered pair (or possibly an ordered n-tuple) of entities one of which is the extension."[19]

Putnam notes that if his analysis of meaning is correct, then the traditional problem of meaning divides into two subproblems: that of accounting for the determination of reference and that of providing criteria for individual competence. Let us consider the second problem first. If follows from Putnam's thesis that the individual need not have a highly-discriminating concept in order to successfully refer--e.g., to refer to elms, I don't

---

[19]*Ibid.*, p. 246.

have to be in a position to distinguish elms from beeches. Yet it is clear that some

discrimination is necessary. If all I know about elms is that they are physical objects, then

we might be inclined to say that my use of the term "elm" fails to secure reference. Putnam

acknowledges that concepts do play a role in the theory of meaning. The problem of

individual competence is to spell out this role.

Let us consider the first problem--that of the determination of extension. It is this

problem which is of interest for our present purposes. In many cases, such as that involving

the beech and the elm, reference is fixed socially, and not individually. This is explained by

Putnam's theory of the division of linguistic labor. Putnam takes this to be an empirical

problem, one to be relegated to sociolinguistics. Solving the problem, Putnam thinks, is

simply a matter of spelling out how the division of linguistic labor works. Putnam takes the

causal theory of reference to fall within this province.

> [T]he fact that, in many contexts, we assign to the tokens of a name that I
> utter whatever referent we assign to the tokens of the same name uttered by
> the person from whom I acquired the name (so that the reference is
> transmitted from speaker to speaker, starting from the speakers who were
> present at the 'naming ceremony,' even though no fixed description is
> transmitted) is simply a special case of social cooperation in the
> determination of reference.[20]

Thus, the position Putnam reaches is that while understanding might be identified

with a narrow psychological state, the consequence is that intension doesn't determine

extension. Putnam does not apply this thought to the problem of skepticism in "The

Meaning of 'Meaning,'" and doesn't draw any conclusions about the relational determination

---

[20]*Ibid.*, p. 246.

of thought content. He clearly wasn't thinking about those problems at the time of writing that paper, and some redeployment is called for to apply the arguments to the problem of skepticism.

This we will take up in the next section when we discuss Putnam's argument in the first chapter of *Reason, Truth and History*. Nonetheless, we can note here quickly the sorts of moves that have to be made. What we need to do is to insist that when we express our thoughts using natural kind terms the contents of the thoughts are to be given by something that does in effect determine the extensions of the terms, whether one thinks of this as an intension of a sort, though not one determined by narrow psychological state, or simply, as Putnam suggests in the paper under consideration, an ordered pair of something determined by narrow psychological state and the actual extension of the term (most philosophers following in Putnam's footsteps have not opted for the latter approach, and indeed it looks as if Putnam drops it as well). In addition, we will want to extend the claim from natural kind terms to all terms we use to express our thoughts about the world around us. We will want to return to the question of whether these moves are justified and what mileage we could get out of them against the skeptic if, indeed, they are. For now I turn to the later development of Putnam's thought about this in the first chapter of *Reason, Truth and History*.

IV. Putnam's Brain in a Vat Argument

Putnam's argument proceeds in two stages. The first stage is intended to establish the correctness of the causal theory of reference (CTR); here we will have to pay attention

to the sorts of considerations Putnam marshals in favor of CTR. The second is supposed to show that if CTR is correct, then the supposition that we are brains in a vat is self-refuting. Given the two stages of his argument, it follows that the supposition that we are brains in a vat is self-refuting.

In the opening paragraph of "Brains in a Vat," Putnam invites us to imagine the following scenario.

> An ant is crawling on a patch of sand. As it crawls, it traces a line in the sand. By pure chance the line that it traces curves and recrosses itself in such a way that it ends up looking like a recognizable caricature of Winston Churchill. Has the ant traced a picture of Winston Churchill, a picture that depicts Churchill?[21]

Putnam's intuition here is that the ant has not succeeded in depicting Winston Churchill-- "The ant, after all, has never seen Churchill, or even a picture of Churchill, and it had no intention of depicting Churchill."[22]

Putnam's intuitions here can be strengthened. Suppose that the ebb and flow of the tides washing up on the beach, together with a slight ocean breeze, just happen to impose marks in the sand which form what we would take to be a recognizable caricature of Winston Churchill. It is obvious that neither the tides, nor the wind, nor any combination thereof, have succeeded in depicting Churchill. The marks on the sand would not be depictions or representations of Churchill. Putnam takes these considerations to support the contention that pictures cannot, in themselves, depict or represent anything at all.

---

[21]H. Putnam, *Reason, Truth and History*, p. 1.

[22]*Ibid.*, p. 1.

Putnam's next move is to extend this observation first to word forms--the token

"WINSTON CHURCHILL" formed in the sand, and then to (what he calls) thought forms--

the contents of representational mental states. Suppose that the ant, rather than tracing what

we would recognize as a caricature of Churchill, had traced what we would recognize as the

name "Winston Churchill." Would the ant have succeeded in referring to Churchill? Of

course not, Putnam says, for we can imagine that the ant has never seen Churchill. So,

Putnam asks, "If lines in the sand . . . cannot 'in themselves' represent anything, then how

is it that thought forms can 'in themselves' represent anything?"[23] Now, although the

randomly formed marks on the sand cannot intrinsically represent Churchill, it is possible

for us to recognize the picture as a caricature of Churchill. The stage setting is now

complete. Putnam is ready to pose the central question of his paper: what is it in virtue of

which we can represent the picture as a caricature of Churchill, while the ant or the wind

cannot? More generally, the question is this: what are the conditions under which we can

refer to and represent something?[24]

One answer here, which Putnam rejects, is that it is in virtue of possessing

intentionality that we can recognize the picture as a caricature of Churchill. Putnam

considers the following argument.

---

[23]*Ibid.*, p. 2.

[24]It is worth noting that the argumentative strategy Putnam is setting up here is Kantian in style. He wants to first argue for a set of necessary conditions under which thought and reference are possible (noting that these are necessary preconditions on giving an argument for skepticism), and then to show that these condition's obtaining precludes the possibility that skepticism is true or, more carefully, that a sound argument for skepticism can be given.

> [W]hat we said about the ant's curve applies to any physical object. No physical object can, in itself, refer to one thing rather than to another; nevertheless, thoughts in the mind obviously do succeed in referring to one thing rather than to another; So thoughts (and hence the mind) are of an essentially different nature than physical objects. Thoughts have the characteristic of intentionality--they can refer to something else; nothing physical has 'intentionality,' save as that intentionality is derivative from some employment of that physical thing by a mind. Or so it is claimed. This is too quick; just postulating mysterious power of mind solves nothing. But the problem is real. How is intentionality, reference, possible?[25]

To advocate such an argument, Putnam thinks, is to commit oneself to a magical theory of reference. It is simply unscientific, Putnam says, to postulate some 'magical connection' between a name and the referent of that name. "What is important to realize," Putnam says, "is that what goes for physical pictures also goes for mental images, and for mental representations in general; mental representations no more have a necessary connection with what they represent than physical representations do. The contrary supposition is a survival of magical thinking" (RTH, p. 3). Putnam's point here is that what needs to be explained is how reference is possible--postulating some magical connection between names and their bearers is no explanation at all.

Putnam advances the causal theory of reference as a *scientific* alternative to the magical theory of reference. The following thought experiment is intended to establish the correctness of the causal theory of reference. Suppose, long ago, a number of humans had been abducted and deposited on some distant planet, Xanx, by alien spacemen. Suppose that generations have persisted and flourished on Xanx. These subsequent generations, although

---

[25]*Ibid.*, p. 2.

quite similar to us, have never seen trees. For we can imagine that on Xanx, there are no trees--perhaps the only vegetable life which exists on Xanx is in the form of molds. Now, imagine that one day an alien space craft accidently deposits a picture of a tree on Xanx. Suppose, further, that this is the only contact Xanxiens have with the aliens. We can imagine the Xanxiens puzzling over the picture--we can imagine them asking: 'What on Xanx is it? A building? A canopy? Perhaps some sort of animal?' Suppose that the Xanxiens never come close to the truth.

As Putnam notes, we would say that the picture is a representation of a tree. But, for the Xanxiens, the picture only represents some strange object whose nature and function is beyond their ken. Let us suppose that one of the Xanxiens forms an image in her mind which is exactly like the image I would form, if I had seen the picture. The question with which Putnam is concerned is this: Is her mental image a representation of a tree? One might be inclined to argue that her mental image really is a representation of a tree, since the picture which caused her mental representation was, itself, a representation of a tree. After all, there is a causal connection (albeit a rather odd one) between actual trees and her mental image.

Putnam concedes the point, only to rebut it by modifying the thought experiment so that this supposed causal connection is entirely absent. We could, for example, suppose that the picture was not, in fact, a picture of a tree, but rather was the accidental result of some spilled paints. In this case, there is no causal connection between actual trees and the Xanxien's mental image. The Xanxien's mental image would not be a representation of a tree, any more than the ant's 'caricature' would be a representation of Winston Churchill.

What is lacking in both cases, Putnam quickly points out, is the appropriate causal connection with objects in the environment--the Xanxien having the appropriate causal connection with trees in the former case, and the ant's having the appropriate causal connection with Churchill in the latter case.[26]

Putnam runs a parallel argument in the case of words. "A discourse on paper might seem to be a perfect description of trees, but if it was produced by monkeys randomly hitting keys on a typewriter for millions of years, then the words do not refer to anything."[27] If, by chance, monkeys had succeeded in typing a copy of *Hamlet*, the words would not thereby refer to anything.[28] This is, Putnam thinks, "a striking demonstration of an important conceptual truth."[29] Namely,

> . . . that even a large and complex system of representations . . . still does not have an intrinsic, built-in, magical connection with what it represents--a connection independent of how it was caused and what the dispositions of the speaker or thinker are. And this is true whether the system of representation . . . is physically realized--the words are written or spoken, and the pictures are physical pictures--or only realized in the mind.[30]

---

[26]As one might expect, a lot hinges on how we are supposed to understand the expression "the appropriate causal connection." We will turn to this shortly.

[27]*Ibid.*, p. 4.

[28]If this were true, it would follow that I would have to rule out the possibility that *Mind, Language and Reality* was produced by monkeys, before I could take be certain that what is said therein is meaningful--a necessary, though not sufficient condition on meaningfulness.

[29]H. Putnam, *Reason, Truth and History*, p. 5. It is interesting that Putnam labels this a "conceptual" truth.

[30]*Ibid.*, p. 5.

From the above considerations, Putnam draws the conclusion that "thought words and mental pictures do not intrinsically represent what they are about."[31] It is a necessary condition on someone's referring to some particular object, that the individual's life history include causal contact with those sorts of objects.

One point that I wish to stress, before turning to Putnam's response to the Brain in a Vat Argument is that the considerations weighing in favor of the causal theory of reference are completely a priori.[32] Indeed, Putnam's investigation begins with the question: what are the conditions under which reference is possible? After giving a number of thought experiments, that is, after engaging in a bit of conceptual analysis, he concludes that it is possible for a subject to refer to some object only if that subject has had causal contact with

---

[31]*Ibid.*, p. 5.

[32]Putnam, himself, claims that the arguments he advances in support of the causal theory of reference are a priori, put he claims they are not a priori "in the old, 'absolute' sense'." They are a priori, Putnam claims:

> . . . in the sense of inquiring into what is reasonably possible assuming certain general premisses, or making certain very broad theoretical assumptions. Such a procedure is neither 'empirical' nor quite 'a priori,' but has elements of both ways of investigating. In spite of my procedure, and its dependence upon assumptions which might be described as 'empirical' (e.g. the assumption that the mind has no access to external things or properties apart from that provided by the senses), my procedure . . . is an investigation . . . of the preconditions of reference and hence thought—preconditions built into the nature of our minds themselves, though not . . . wholly independent of empirical assumptions. (p. 16).

This claim, I think, is unfortunate. The Brain in a Vat Argument issues a challenge to the possibility of our having any knowledge of the external world—i.e., to the possibility of our having any empirical knowledge. To make any empirical assumptions in responding to the Brain in a Vat Argument would be to presuppose that we already have some knowledge about the external world, which is to beg the question against the advocate of the Brain in a Vat Argument.

those kinds of objects. The point I wish to emphasize is that nothing Putnam has said even purports to establish the claim that we ever actually succeed in referring to anything. As noted above, Putnam's stated purpose (in the first stage of his argument) is to elucidate the "preconditions for thinking about, representing, [and] referring to."[33] The causal theory of reference is intended to state just what those preconditions are.

This way of understanding the first stage of Putnam's argument has a cost--we could no longer run the argument from magic, for there would be no basis for saying that CTR is a *scientific* alternative to the magical theory of reference. However, this concession is necessary, if CTR is to have any force against the skeptic. Recall that a satisfactory response to skepticism cannot appeal to any facts about the world.

The second stage of Putnam's argument is designed to show that if we really are (and have always been) brains in a vat, those preconditions could not be met. Before going into Putnam's reasoning, it will be useful to describe, more fully, the version of the Brain in a Vat Argument that Putnam takes up. The version of the challenge that Putnam takes up is more specific than the way the challenge is typically advanced. (This has important ramifications, as I will point out later.)

The typical version of the Brain in a Vat Argument runs as follows. Imagine an individual who (unbeknownst to him) has been subjected to an operation in which an evil scientist removes the subject's brain from his body and places it in a vat of nutrients. Suppose that the subject's nerve endings are connected to a highly sophisticated computer system. By running the appropriate program, the scientist is able to obliterate all of the

---

[33]*Ibid.*, p. 16.

subject's memories of the brain operation. The result is that from a purely phenomenal standpoint, the subject's life seems to proceed without disruption. Nonetheless, his life would be a mere illusion. By running any number of programs, the scientist could systematically manipulate the subject's experiences. In such cases, none of his experiences would be veridical. The question, then, is how do we know we are not in this predicament?

Now, as Putnam notes, the purpose of discussing the brain in a vat possibility is to raise traditional questions about our knowledge of the external world in a modern way. Putnam's concern, however, is to raise a question about the mind/world relationship. He sets out the predicament as follows.

> Instead of having just one brain in a vat, we could imagine that all human beings . . . are brains in a vat. . . . Perhaps there is no evil scientist, perhaps . . . the universe just happens to consist of automatic machinery tending a vat full of brains and nervous systems.
> This time let us suppose that the automatic machinery is programmed to give us all a collective hallucination, rather than a number of separate unrelated hallucinations.[34]

What is significantly different about Putnam's formulation of the brain in a vat predicament is the supposition that the entire universe consists of automatic machinery tending the vat. In fact, Putnam goes on to suppose that "the automatic machinery is supposed to come into existence by some kind of cosmic chance or coincidence (or, perhaps, to have always existed)."[35] Moreover, Putnam supposes that "all sentient beings (however minimal their

---

[34]*Ibid.*, pp. 6-7.

[35]*Ibid.*, p. 12.

sentience) are inside the vat."[36]   In effect, we are asked to suppose that no sentient being has ever existed outside the vat.

Putnam invites us to suppose that this whole story is actually true. The question with which Putnam is concerned is this: "Could we, if we were brains in a vat in this way, say or think that we were?"[37] Putnam argues that if we really are brains in a vat (in the above sense), it would not be possible for us to say or think that we are brains in a vat.

Putnam's stated purpose in the second stage of his argument is to show that "the supposition that we are actually brains in a vat [in this way], although it violates no physical law, and is perfectly consistent with everything we have experienced, cannot possibly be true. It cannot possibly be true, because it is, in a certain way, self-refuting."[38] Put in terms of possible worlds, there is a possible world in which all sentient beings are brains in a vat. "The humans in that possible world," Putnam says, "have exactly the same experiences that we do. They think the same thoughts we do (at least, the same words, images, thought-forms, etc., go through their minds.)"

Nonetheless, Putnam thinks that we can know that we are not brains in a vat (in the above sense). Putnam's argument is supposed to show that the actual world could not possibly be a world in which all sentient beings are (and have always been) brains in a vat. Putnam gives the following sketch of his argumentative strategy.

---

[36]*Ibid.*, p. 12.

[37]*Ibid.*, p. 7.

[38]*Ibid.*, p. 17.

[A]lthough the people in that possible world [the world in which all sentient beings are (and have always been) brains in a vat] can think and 'say' any words we can think and say, they cannot . . . refer to what we can refer to. In particular, they cannot think or say that they are brains in a vat (even by thinking 'we are brains in a vat').[39]

The reason why they cannot think or say that they are brains in vat, even by thinking "we are brains in a vat" is that they lack the appropriate causal connection with actual brains and actual vats.

More generally, they cannot refer to external objects at all since, by hypothesis, they have not had any causal contact with external objects whatsoever. For example, "when the brain in a vat (in the world where every sentient being is and always was a brain in a vat) thinks 'there is a tree in front of me,' his thought does not refer to actual trees."[40] If his thought does not refer to actual trees, what then does it refer to? Putnam is willing to entertain a variety of answers. He considers the following:

On some theories that we shall discuss it [the brain in a vat's thought] might refer to trees in the image, or to the electronic impulses that cause tree experiences, or to the features of the program that are responsible for those electronic impulses. These theories are not ruled out by what was just said, for there is a close causal connection between the use of the word 'tree' in vat-english and the presence of trees in the image, the presence of electronic

---

[39]*Ibid.*, p. 7.

[40]*Ibid.*, p. 14. Putnam's talk of thoughts referring to objects is non-standard. We typically say that our thoughts either represent or fail to represent some state of affairs, or that our thoughts are about this or that. In contrast, we typically talk of words (or other such linguistic entities) as referring or failing to refer to objects. Although Putnam's talk of thoughts referring to objects is non-standard, I do not think that anything crucial hinges on this way of talking. For example, where Putnam says "his thought does not refer to actual trees," we might say that his thought is not about trees. For ease of exposition, I will follow Putnam's way of talking.

impulses of a certain kind, and the presence of certain features in the machine's program.[41]

On these theories, Putnam notes, the brain in a vat is correct in thinking that there is a tree in front of it, given what "tree" and "in front of" refer to in Vat-English. The truth conditions for 'there is a tree in front of me' in Vat-English "are simply that a tree in the image be 'in front of' the 'me' in question--in the image--or, perhaps, that the kind of electronic impulse that normally produces this experience be coming from the automatic machinery, or perhaps, that the feature of the machinery that is supposed to produce the 'tree in front of one' experience be operating."[42] These conditions, Putnam claims, are certainly satisfied.

By the same line of reasoning Putnam concludes that in Vat-English 'vat' refers to vat in the image. From this observation, Putnam concludes:

> It follows that if their 'possible world' is really the actual one, and we are really the brains in a vat, then what we now mean by 'we are brains in a vat' is that we are brains in a vat in the image or something of that kind (if we mean anything at all). But part of the hypothesis that we are brains in a vat is that we aren't brains in a vat in the image (i.e. what we are 'hallucinating' isn't that we are brains in a vat). So, if we are brains in a vat, then the sentence 'We are brains in a vat' says something false (if it says anything). In short, if we are brains in a vat, then 'We are brains in a vat' is false. So it is (necessarily) false.[43]

Putnam's claim, then, is that if we can entertain the supposition that we are brains in a vat, then the sentence "Possibly, we are brains in a vat" is necessarily false.

---

[41]*Ibid.*, p. 14.

[42]*Ibid.*, p. 14.

[43]*Ibid.*, p. 15.

In the next section, we shall take up an objection to Putnam's argument, the thrust of which is that Putnam's argument is incoherent.[44] In section VI, I propose a reconstruction which aims to avoid this charge. However, I shall argue in section VII that even this more refined reconstruction has fatal flaws.

V. An Objection to Putnam's Brain in a Vat Argument

The thrust of the objection, I take it, centers on Putnam's claim that if we really are brains in a vat, then the sentence "We are brains in a vat" is false. Putnam's claim seems to have the following form: if p, then 'p' is false, where both instances of 'p' are replaced by sentences of the same language. But, if 'p' is false it follows that not-p. So, it looks as though the form of Putnam's argument reduces to: if p, then not-p, hence, not-p. This, however, cannot correctly capture the form of Putnam's argument, it is argued,

> . . . because in the conditional that Putnam affirms the truth conditions of the antecedent are the ones it has when interpreted in English, while the reason the sentence quoted in the consequent is false is that its truth conditions are given relative to Vat-English. So the sentence quoted in the consequent is not interpreted in the same way as the one used as the antecedent. From the falsity of the sentence quoted in the consequent, interpreted relative to Vat-English, we can conclude nothing directly about the truth or falsity of the antecedent, interpreted relative to English. But what we want to know is whether the sentence interpreted relative to English is true or false.[45]

---

[44]Kirk Ludwig, "Brains in a Vat, Subjectivity, and the Causal Theory of Reference," *Journal of Philosophical Research*, 1992, Vol. XVII, pp. 313-345.

[45]*Ibid.*, pp. 332-333.

But, if we interpret the conditional relative to English, while interpreting the sentence quoted in the consequent relative to Vat-English, we end up with an incoherent argument. For the argument would then run as follows:

> If the English sentence 'We are brains in a vat' is true . . . then we speak not English, but Vat-English. Since in vat-English 'brains' and 'vats' do not refer to brains and vats, but at best, perhaps, to elements in the machine or image generated by the machine, in Vat-English 'We are brains in a vat' is false. That is to say, if our sentence 'We are brains in a vat' is true, then it is, after all, false. Therefore, our sentence 'We are brains in a vat' is false, and necessarily so.[46]

The difficulty here is that if we attempt to say "If the sentence 'We are brains in a vat' is true, then it is false" we are shifting between two different languages. And, we can say that a sentence is true (or that a sentence is false) only relative to some specified language.

To clean up Putnam's incoherence, Ludwig suggests the following reconstruction of the BIV argument.[47]

[1]   I am wondering whether I am a brain in a vat. (Introspection)

[2]   Necessarily, for any x, if x is a brain in a vat, x is not wondering whether x is a brain in a vat. (CTR)

[3]   Therefore, I am not a brain in a vat. (from 1 and 2)

The difficulty with this argument, according to Ludwig, focuses on premise [1]. Recall that it is part of the description of Putnam's thought experiment that the subjective character of

---

[46]*Ibid.*, p. 333.

[47]*Ibid.*, p. 333.

my thoughts (i.e., how things seem to me through introspection) is qualitatively indistinguishable from that of my envatted counterpart. The subjective character of our thoughts is the same; however, the contents of our beliefs and the meanings of our terms differ.

For example, when I think *there is a tree in my front yard* the "images floating in my head" (so to speak) are introspectively indistinguishable from the images floating in my envatted counterparts head when she thinks *there is a tree in my front yard.* Yet my belief is about a tree, whereas my counterpart's belief is about the electronic impulses that cause tree experiences.

A difficulty emerges, according to Ludwig, when we consider that the only way for me to know that I am having a certain thought rather than another is the introspective quality or the subjective character of that thought. If this is correct, then the introspective quality of my thought is not sufficient to determine whether, for example, my belief is about an actual tree or whether it is about an electronic impulse. By the same token, when I think *I am wondering whether I am a brain in a vat*, I don't know whether I am entertaining a thought about actual brains and actual vats or whether I am entertaining thoughts about electronic impulses.

The charge, then, is that Putnam is committed to the claim that we do not have knowledge of the contents of our own thoughts. The argument, as Ludwig sets it out, goes as follows:

> [I]f we have knowledge of thought contents, then the subjective character of our conscious thoughts determines that knowledge. Knowledge of thought

> contents determines what the thought contents are. By transitivity, subjective
> character determines thought contents. So if we have knowledge of thought
> contents, they are determined by subjective character. Since Putnam denies
> that thought contents are determined by subjective character, if we accept that
> knowledge of thought contents rests on subjective character, we must see
> Putnam as committed to holding that we do not have knowledge of thought
> contents.[48]

I shall not pause here to evaluate this criticism, for I think there is a way of reconstructing

Putnam's argument which avoids the charge.

VI. An Alternative Reconstruction of Putnam's BIV Argument

Now, I should say at the outset that I do not think it is possible to show that we are

not, in fact, brains in a vat. I do, however, think that it is possible to give an argument,

granting CTR, which shows that the skeptic cannot give a good argument which shows that

it is possible that we are (and have always been) brains in a vat. The reconstruction I propose

is designed to show that (given the causal theory of reference), any argument which purports

to show that we are (and have always been) brains in a vat is self-refuting.

Before giving the argument, there are two significant differences between the

argument Putnam gives and the reconstruction I propose. First, Putnam attempts to show

that the *supposition* that we are (and have always been) brains in a vat is self-refuting. In

contrast, my argument is designed to show that any *argument* given in support of the claim

that we cannot know that we are not (and have not always been) brains in a vat is self-

refuting. The second difference is that my conclusion is somewhat weaker than the

---

[48]*Ibid.*, p. 334.

conclusion Putnam draws. I do not follow Putnam in arguing that the actual world is not a world in which we are (and have always been) brains in a vat. Rather, I argue that the skeptic has failed to show that it is epistemically possible that we are (and have always been) brains in a vat.

The argument goes as follows:

[1]   To give an argument for the claim that we cannot know that we are not brains in vats, we must be able to entertain the supposition that we are brains in vats.

This follows from the fact that we must use the relevant concepts in the premises and the conclusion of the argument.

[2]   To entertain the supposition that we are brains in vats, we must not be brains in vats.

This follows from the Causal Theory of Reference.

[3]   Therefore, to given an argument for the claim that we cannot know that we are not brains in vats, we must not be brains in vats.

Now, it is a reasonable constraint to put on an argument which aims to show that we don't know something that being in a position to give it, that is, to know or entertain its premises, not validate what we are being said to fail to be able to know. For if we know this, then we know that if we are in a position to put the argument forward, then, on the assumption that

we know we are in that position, we are in a position to come to know the thing we are said

not to be in a position to know. We will call such an argument self-refuting (SR)

> (SR)   A skeptical argument for the claim that we cannot know that p is self-refuting
> if to be able to knowingly put it forward we must be in an epistemic position
> to establish p.

It is clear that if CTR is correct, then [3] follows immediately. So, if we can know CTR on

a priori grounds, as needed, it follows that we can know [3]. And if we can know that we

are in a position to put forward an argument for the claim that we cannot know we are brains

in a vat, then from that and [3] we are put in a position to infer that we are not brains in vats.

Thus, any skeptical argument for the claim that we cannot know we are brains in a vat is self-

refuting, if CTR can be established *a priori*. Thus, we can establish the conclusion that

skeptical arguments of this sort are self-refuting without having to enter as a premise that we

know that we entertain suppositions about brains in vats.

However, it might be objected that to know [3] would require us to know we were

not brains in vats, for otherwise we would not be thinking the thoughts we are with those

words. But the argument can proceed with a place-holder in the position of the sentences

expressing what we are said not to know. The point is that to put forward an argument

knowingly, the skeptic must assume we can know the premises of it, and if he can know

those, then he can know [3] as well, given the *a priori* knowability of CTR, and he is in a

position to establish what the argument says he cannot establish. To get our result, we do not

have to assume that we can show we are not brains in vats, but only that if we could given

an argument to show that we cannot know we are not brains in vats, then we could know we are not brains in vats.

VIII. Putnam's Retreat on Metaphysical Necessity

Thus far I have been clarifying the force of an argument based on CTR against skepticism. I have argued that Putnam need to argue for CTR on an *a priori* basis, and that despite his sometimes suggesting that it is a scientific alternative to the magical theories of reference and representation, his actual procedure suggests he is thinking of it as an *a priori* argument. In the final section of this chapter I shall return to the question of whether the thought experiments really are entirely free from empirical assumptions. But in this section I wish to take up a development in Putnam's views about the status of CTR. For in later work, particularly in "Is it Necessary that Water is H_2O?", Putnam questions whether it makes sense to talk of water's being H_2O in every possible world. I will review the reasons Putnam offers for this position, and then assess both its implications, and the adequacy of Putnam's reasoning for it.

A.J. Ayer attributed to Putnam the view that it is inconceivable that water is not H_2O, and on that basis objected that Putnam's view is clearly mistaken. Putnam's response was that he had never maintained that it was inconceivable that water is not H_2O, but only that it was impossible that water is not H_2O. It was his reflection on this response that led Putnam to revise his views. By the time he wrote "Is it Necessary that Water is H_2O?", he came to realize that his response to Ayer was "less than just."

The standard argument for the necessity of the claim that water is H_2O has it that substance terms, like proper names or descriptions, can be used either rigidly or non-rigidly. "Thus, if I were to use the word 'water' to refer to whatever stuff has certain observable characteristics in the hypothetical situation (regardless of its chemical composition) Kripke would say that I was using the term 'non-rigidly'." But, Putnam continues, "If I were to use it to refer to whatever stuff is substance-identical with the stuff that has those observable characteristics in the actual world, whether or not it has them in the hypothetical situation, then Kripke would say I was using the term rigidly."[49] In this case, if the stuff that has those observable characteristics in the actual world is substance Y, then (so the argument goes), water is Y in every possible world.

What determines what is the same substances as a given sample? Putnam's considered view is that

[SS]   X is the same substance as Y iff X has the same physiochemical composition as Y and X obeys the same laws as Y.[50]

Then we say that

[WI]   X is water iff X is substance-identical "with (most of) the paradigms in our actual environment (limited both to the actual world and to the available part of the actual world)."[51]

---

[49]Putnam, "Is it Necessary that Water is H_2O?" in *The Philosophy of A.J. Ayer*, L.E. Hahn, ed., (La Salle: Open Court, 1992), p. 437.

[50]*Ibid.*, p. 439.

[51]*Ibid.*, p. 439.

It is this last formulation that can help us to see what Putnam's later worries come to. What Putnam is concerned about is the fact that this apparently makes "logical possibilities . . . dependent on empirical facts."[52] If I understand his line of reasoning, it is the following. A logical or metaphysical necessity is supposed to be one that holds in every possible world. But the supposed logical or metaphysical necessity of water's being H_2O is not independent of what contingent facts obtain, namely, the contingent facts about what stuff we refer to using the term "water." But nothing which is dependent on a contingent fact can itself be a necessary truth, for that is to say that necessity is contingent, and that is incoherent.

Thus, Putnam concedes, 'I now think that the question "what is the necessary and sufficient conditions for being water in all possible worlds?' makes no sense at all. And this means that I now reject 'metaphysical necessity'."[53] In other words, Putnam is saying that if things had been different, what water would have been constitutively would have been different. Thus it can't be that water must be H_2O in all possible worlds, for then this last claim would make no sense.

I am by no means assured that Putnam's reasoning here is correct. However, before I turn to a critical evaluation of Putnam's reasoning, I want to point out what the consequences are for the use of CTR against skepticism, at least insofar as it is based on the rigid use of kind terms, if his reasoning is correct. Let us, then, assume, for the moment, that Putnam's reasoning is correct.

---

[52]*Ibid.*, p. 437.

[53]*Ibid.*, p. 437.

If Putnam intends for CTR to be reinterpreted in the way he reinterprets his thesis about the identity of water and H_2O, then it follows that CTR is not necessarily true--i.e., not true in all possible worlds. However, if this is the case, then we would not be in a position to argue that a necessary condition on entertaining thoughts about brains and vats is that one not be a brain in a vat. This undermines both the stronger and the weaker arguments against skepticism that we considered on the basis of CTR. It was the claim that causal interaction was necessary for reference that allowed us to provide an *a priori* response to the skeptic, that is, one that didn't beg the question. If all we know is that if the world is a certain way, then CTR is true, then we have no non-question begging response to skepticism.

Let us now turn to the question of whether Putnam's reasoning is correct. There is, I think, one way to circumvent Putnam's conclusion.[54] However, even if this strategy is pursued, it will be of no anti-skeptical value. There is a way of understanding how natural kind terms work that would allow us to say that if "water" has been applied to the stuff which is predominately H_2O, then necessarily water is H_2O.

It is that [WI] serves as a way of specifying the meaning of "water", that is, that it is to be interpreted as claiming that "water" expresses whatever concept is expressed by a predicate in some language which picks out the property of being a substance identical "with (most of) the paradigms [of the referent of "water"] in our actual environment (limited both to the actual world and to the available part of the actual world)." This would then allow us to say that *if* that substances is H_2O, then "water" means "H_2O." And if that is true, then it

---

[54] I am indebted to Kirk Ludwig for the following points.

follows of course that necessarily water is H_2O, where the necessity concerned is conceptual necessity.

Unfortunately, this is no help. We use the term "water" in characterizing our own and others' beliefs. We have two choices on the view being entertained. Either we falsely attribute such beliefs to ourselves because we don't know what concept is expressed by "water" and so don't have the beliefs we attribute (or at best have them only accidently relative to the attribution), or we correctly attribute such beliefs to ourselves on those occasions but don't know what their contents are. In either case, we would not be in a position to draw any inference from our beliefs to the nature of our environments. I conclude, then, that CTR, pursued on the basis of reflections on natural kind terms, is a blind alley as far as skepticism is concerned.

IX. Are there Empirical Presuppositions to Putnam's Thought Experiments?

The previous section was primarily concerned with developments in Putnam's views on the modal strength of claims about what natural kinds are. The application of CTR depends on supposing that CTR is to be supported by the sorts of thought experiments given in "The Meaning of 'Meaning'." Putnam clearly seems to think the same sort of theory is being advanced in each place. Yet, in *Reason, Truth and History*, Putnam offers thought experiments in support of this view which are *prima facie* independent. Therefore, it will be important to ask whether they are successful if we take them as independent of considerations about natural kind term semantics.

I wish to urge two points. The first is that it is not clear that intuitions elicited in the thought experiments are independent of empirical presuppositions. Indeed, Putnam himself suggests that they are not, and I think for good reason. It is not accidental that he rejects an anti-reductionist view as *unscientific* before introducing the thought experiments. For the intuitions elicited are really elicited in the light of the *empirical methodological presupposition* that there is some scientific reductive account to be had, some robust explanatory story, about why we are able to represent what we are able to represent.

For clearly if we assume that there is, and ask under what conditions someone is able to represent something, we will consider what sort of systematic account might be forthcoming which will be broadly reductive in character, and it seems pretty obvious that any such account which is not completely arbitrary and which is accepted by our scientific world view is going to have to treat at least as a necessary condition on representation (in some sense of "necessary") causal contact with things of the general sort being represented. I submit that in fact this is the source of the intuitions Putnam elicits, and that therefore they are intuitions not just about what is possible, but about what resources would have to be appealed to is we were to provide a scientific reductive account of representation.

But this is no use against skepticism, because the presuppositions are themselves, as Putnam admits, quasi-empirical in character. Thus, independently of the considerations of the previous section, we have reason to think that the thought experiments proposed in the first chapter of *Reason, Truth and History* are unsuccessful in establishing a thesis which is of use against the skeptic. We could use the results only if we know that their empirical

presuppositions were correct, but it falls in the scope of the skeptic's argument, and the force of the conclusion is therefore rendered nugatory.

The second point I wish to urge is that, independently of the first, the charge that we leveled against Burge's thought experiments in support of perceptual externalism can also be lodged against Putnam's thought experiments. For what Putnam does is to ask whether we think that in a spatial world in which we are in causal interaction with material objects we think our thought contents would vary with what objects we were in causal contact with.

Now as we observed about Burge's similar argument, this does not show that having the relevant representation contents requires being in a spatial world, but at most that if we have them in a spatial world, they will vary with what causes them. So, at best, we have a conditional conclusion, if we are in a spatial world, then . . . . But then to use this against the skeptic, we must discharged the antecedent, but that falls within the scope of the skeptic's conclusion, and so to do so would be to beg the question against the skeptic.

CHAPTER V
DAVIDSON'S INTERPRETATION-BASED ARGUMENTS

I: Introduction

Descartes began his project of reconstructing the edifice of knowledge from within the sanctity of the first-person perspective. "I am, I exist" was his Archimedean point. He sought to restore (much of) our knowledge of the world around us on the basis of this single proposition. A suggestion, implicit in the work of Donald Davidson, is that we need not take the Cartesian perspective of solitary self-examination as our methodological starting point. Instead, Davidson urges that we adopt the stance of the radical interpreter as our fundamental starting point for philosophical inquiry. From this vantage point, Davidson claims to have established that radical skepticism about the world is unintelligible. The aim of this chapter is to clarify and evaluate Davidson's arguments for this claim.

This chapter is divided into six sections. In the next section, I will introduce the notion of a radical interpreter. In section III, we turn directly to Davidson's much celebrated Omniscient Interpreter Argument. This argument provides a principled reason for thinking that our beliefs about the world are largely true. From this conclusion, it is only a small step to showing that radical skepticism is misguided. Our evaluation of this argument will carry over into section III, where we will have to discuss some of the general features of Davidson's method of interpretation in connection with his externalism. In section V, I argue that Davidson's Omniscient Interpreter Argument tacitly appeals to empirical

presuppositions. In section VI, I examine some of Davidson's more recent work. In particular, I take his Argument from Objectivity.

II. Davidson's Method of Interpretation

What, then, is the "stance" of the radical interpreter? Much of Davidson's work can be viewed as an answer to a single question: what are the conditions under which communication (or interpretation) is possible? To answer this question it shall be useful, though not necessary, to imagine oneself thrown into a jungle and forced to construct a theory of interpretation for the local inhabitants. Your situation is this: you have no idea what the sentences they utter mean or what the contents of their beliefs and other attitudes are.

What must you presume, in constructing your theory? You are forced to regard the natives as engaging in intentional discourse; that is, you are forced to assume that they are, in fact, speaking a language. You must therefore assume that your new found fellows are rational agents. You have no choice but to impose your standards of logic upon them. Your job is to construct a theory which maximizes the intelligibility of their utterances.

How, then, should you proceed? Routine observations would reveal which sentences a given native accepts as true under which circumstances. With this information at your disposal, your job is to determine the meanings of the sentences he utters and the contents of his thoughts. "The methodological problem of interpretation," Davidson says, "is to see

how, given the sentences a man accepts as true under given circumstances, to work out what his beliefs are and what his words mean."[1]

This methodological problem is best resolved, Davidson thinks, by constructing a Tarski-style theory of truth for the speaker's language. The assignment of truth conditions to the speaker's utterances is informed by the speaker's assent to sentences held true under recognized conditions. In your role as radical interpreter, you have no choice but to infer meaning, and hence belief, from the events and objects in the world which you take to have prompted the speaker's assent. "[T]he interpreter," Davidson says, " . . . must take into account the causal interaction between world and speaker in order to find out what the speaker means, and hence what he believes."[2]

Davidson thinks that his method of interpretation delivers a decisive verdict against the skeptic. In a frequently quoted passage, he says:

> What stands in the way of global skepticism of the senses is, in my view, the fact that we must, in the plainest and methodologically most basic cases, take the objects of a belief to be the causes of that belief. And what we, as interpreters, must take them to be is what they in fact are. Communication begins where causes converge: your utterance means what mine does if belief in its truth is systematically caused by the same events and objects.[3]

At first glance, one might think that this method can be successfully employed only if we first know that skepticism about the external world is false. After all, if successful communication depends on our being in a position to identify the objects and events in a

---

[1] Donald Davidson, "Thought and Talk," p. 162.

[2] Donald Davidson, "Empirical Content," p. 332.

[3] Davidson, "A Coherence Theory of Truth and Knowledge," p. 317.

speaker's environment which prompt assent to sentences, then we obviously can't assume that we ever correctly interpret another speaker. More obvious still, if we are giving an argument against skepticism about the external world, then we cannot assume that there are other speakers--for the existence of other people falls within the scope of the skeptic's argument. These observations would seem to suggest that Davidson's methodology presupposes, and hence cannot be employed to establish, the falsity of skepticism about the external world.

I think that this line of criticism misses the point of Davidson's discussion. We should take Davidson to be giving an analysis of the conditions under which communication (or interpretation) is possible. We should view him as engaging in a bit of conceptual analysis. Indeed, in introducing the notion of a radical interpreter, I asked you to *imagine* yourself thrown into a jungle. This is of a piece with traditional philosophical reflection.

What falls out of Davidson's *a priori* reasoning is that we cannot fix the contents of a person's beliefs independently of what caused them. In his own words:

> What I take to be the important aspect of this approach is apt to be missed because the approach reverses our natural way of thinking of communication derived from situations in which understanding has already been secured. Once understanding has been secured we are able, often, to learn what a person believes quite independently of what caused him to believe it. This may lead us to the crucial, indeed fatal, conclusion that we can in general fix what someone means independently of what he believes and independently of what caused the belief. But if I am right, we can't in general first identify beliefs and meanings and then ask what caused them. The causality plays an indispensable role in determining the content of what we say and believe.[4]

---

[4] *Ibid.*, p. 317.

Davidson contends that given the way beliefs are identified and described, it is not possible for us as interpreters to discover massive error in another--hence the conclusion that global skepticism about the external world is unintelligible.

What I have just said is that we cannot fix what another's beliefs and meanings are without attention to what determines which beliefs that person holds true directed toward sentences. It might look as if this is just an epistemic point, a point about how we have access to, gain knowledge of, the beliefs of others, and not about what logically determines what their beliefs are. If this were the extent of the point, however, it would not be much help. But Davidson's point is that because the stance of the interpreter is methodologically basic in understanding communication, since language is a tool above all for enabling different thinkers to communicate with one another in a public medium, what is epistemically necessary in interpretation reveals conceptually necessary conditions on what emerges from the procedure. That is, what the interpreter must assume to interpret a speaker is constitutive of his subject matter.

Davidson's initial argument from the standpoint of the radical interpreter to the claim that massive error is unintelligible is found in "The Method of Truth in Metaphysics." There, he asks the question: Why must communication rest on a largely correct, shared view of reality? This is one of the central questions of his paper. He tackles this question in two stages. In the first stage, he aims to show that communication depends on a shared view of reality; and in the second stage, he aims to show that this shared view of reality must be largely correct.

Let us consider each of these stages in turn. His argument for the first stage is contained in the following paragraph.

> [W]e damage the intelligibility of our readings of the utterances of others when our method of reading puts others into what we take to be broad error. We can make sense of differences all right, but only against a background of shared belief. What is shared does not in general call for comment; it is too dull, trite, or familiar to stand notice. But without a vast common ground, there is no place for disputants to have their quarrel.[5]

In other words, we can communicate with others only if we see them as holding beliefs about the world which are similar to our own. Davidson recognizes that there are bound to be subtle differences in our belief systems. His point is that these differences can be rendered intelligible only against a background of shared beliefs.

Take, for example, my belief that Florida will crush Florida State next year. Someone may disagree with me, but only if we agree on a lot of other things: that the two schools have football teams, that the two teams will play each other again next year, that football is a sport, that such and such a score constitutes a crushing victory in a football game, and so forth. In short, I can make sense of another's disagreement, but only if I suppose that her beliefs are, on the whole, quite similar to my own. If I should discover that her conception of a football game was that of a poetry recital, what would be called for is not an attribution of a belief about next year's Florida/Florida State game, but a complete revision of my original ascription. Too much disagreement jeopardizes the intelligibility of

---

[5] Davidson, "The Method of Truth in Metaphysics," p. 200.

my ascription. "No wonder," Davidson concludes, "I can interpret your words correctly only by interpreting so as to put us largely in agreement."[6]

This example supports Davidson's contention that communication depends on a shared view of reality, that is, on shared beliefs. Let us turn now to the second stage of his argument. Why does Davidson claim that this shared view of reality is largely correct? His argument is rather brief. He says, ". . . objective error can occur only in a setting of largely true belief. Agreement does not make for truth, but much of what is agreed must be true if some of what is agreed is false."[7] Davidson's claim, then, is that we can discover error in another speaker, but only against a background of largely true belief. In other words, the ascription of false beliefs carries with it the implication that the speaker's belief system is largely consistent and correct. Hence, it is unintelligible to suppose that other speakers could be massively mistaken about the world.

Let us sum up the state of the argument thus far. I contend that Davidson's considerations about interpretation are best construed as an analysis of the conditions under which interpretation is possible. His argument aims to show that we can successfully interpret other speakers only if we *suppose* that their beliefs are largely true. It would be a mistake to think that Davidson relies on the view that we ever successfully interpret another. This would be an empirical claim, given the preconditions he has set down on successful interpretation. To rely on it here would undermine any anti-skeptical force that his method of interpretation may deliver.

---

[6] Davidson, "The Method of Truth in Metaphysics," p. 200.

[7] Davidson, "The Method of Truth in Metaphysics," p. 200.

III: Davidson's Omniscient Interpreter Argument

With this said, let us now turn to Davidson's direct assault on the skeptic. His

argument begins with the claim that to correctly interpret others, we must view them as

holding beliefs which are mostly true. As we have already seen, this claim does not entail

that the beliefs of those we interpret are, in fact, true. To bridge the gap between what is held

to be true and what is true, Davidson introduces the notion of an omniscient interpreter.

> [I]magine for a moment an interpreter who is omniscient about the world, and
> about what does and would cause a speaker to assent to any sentence in his
> (potentially unlimited) repertoire. The omniscient interpreter, using the
> same method as the fallible interpreter, finds the fallible speaker largely consistent
> and correct. By his own standards, of course, but since these are objectively
> correct, the fallible speaker is seen to be largely correct and consistent by
> objective standards. We may also, if we want, let the omniscient interpreter
> turn his attention to the fallible interpreter of the fallible speaker. It turns out
> that the fallible interpreter can be wrong about some things, but not in
> general; and so he cannot share universal error with the agent he is
> interpreting. Once we agree to the general method of interpretation I have
> sketched, it becomes impossible correctly to hold that anyone could be mostly
> wrong about how things are.[8]

On the face of it, the argument here is straightforward. The first premise holds that it is

possible that there is an omniscient interpreter, a being that is "omniscient about the world."

The second premise states that by using "the same method as the fallible interpreter," the

omniscient interpreter "finds" us (fallible speakers) "to be largely consistent and correct."

Since, by hypothesis, the omniscient interpreter has only true beliefs--that is, since he is

---

[8] Donald Davidson, "A Coherence Theory of Truth and Knowledge," p. 317.

omniscient about the world--and since he interprets us such that we are largely in agreement with him, it follows that a significant portion of our beliefs about the external world are true.

The conclusion of this argument, if correct, would be sufficient to blunt the force of radical skepticism. For although we would not be in a position to say of any particular belief that it is, in fact, true, we would be in a position to say of a given belief that it is likely to be true. It is our knowledge that the set of beliefs as a whole is largely consistent and correct which entitles us to confer a favorable epistemic status on any particular belief. What this argument yields, then, is a principled reason for thinking that a significant portion of our beliefs are true.

Given this, we can test individual beliefs by appeal to how well they are supported by the other beliefs we have, for a belief which coheres well with most of what we believe is likely to be true, given that most of our beliefs are true. Thus, as Davidson has put it, belief makes for truth, and thus coherence with belief makes for truth as well. If Davidson is right, this is the truth in the coherence theory of knowledge.

In what follows, I will present a number of formulations of the argument, before settling on the one that I think is the strongest. Richard Foley and Richard Fumerton have suggested what is, perhaps, the most obvious way of reconstructing Davidson's argument.[9]

[1] Possibly, there exists an omniscient interpreter--a being whose beliefs about the world are all true. (premise)

---

[9] This formulation is suggested by Richard Foley and Richard Fumerton in "Davidson's Theism?" pp. 83-89. I say "suggest" because they do not lay out Davidson's argument explicitly. However, the objection they raise against Davidson (discussed below) shows that they must have this sort of formalization in mind.

[2] For any x and for any y, x can interpret y (correctly) only if x's beliefs are largely in agreement with y's. (premise)

[3] If there were an omniscient interpreter, he could interpret us. (premise)

[4] If there were an omniscient interpreter, his beliefs would be largely consistent with our beliefs. (from 2 and 3)

[5] Therefore, our beliefs about the world are largely true. (from 1 and 4)

They object to this argument, noting that to get a valid argument we need to affirm the antecedent of the conditional in premise [3].[10] They remark:

> From (1) If there were an omniscient interpreter employing Davidson's methods of interpretation he would believe that most of what Jones believes is true, how is it supposed to follow that most of what Jones believes is true? From (1) we *can* infer (2) If there were an omniscient interpreter of Jones employing Davidson's methods, most of what Jones believes would be true. But surely we need to affirm the antecedent of this conditional if we are to conclude that most of Jones's beliefs *are* true. That is, we need to affirm that there is an omniscient interpreter of Jones.[11]

I think this is an uncharitable reconstruction of Davidson's argument. The problem is that their formulation (uncharitably) sticks Davidson with a claim that falls within the scope of the skeptic's argument. To affirm the antecedent in [3], Davidson must assert that there is an omniscient interpreter. This is an empirical claim. Hence, it is one to which Davidson cannot appeal in arguing against the skeptic.

Charity demands a better formulation. The following formulation avoids the above difficulty.

---

[10] *Ibid.*, p. 84.

[11] *Ibid.*, p. 84.

[1]   Possibly, there exists an omniscient interpreter--a being whose beliefs about the world are all true. (premise)

[2]   For any x and for any y, x can interpret y (correctly) only if x's beliefs are largely in agreement with y's. (premise)

[3]   Our beliefs are such that if there were an omniscient interpreter, he could interpret us. (premise)

[4]   Therefore, our beliefs are largely in agreement with the beliefs that an omniscient interpreter would have, were one to exist. (from 2 and 3)

[5]   Therefore, our beliefs are largely true. (from 1 and 4)

This formulation is superior to the first in that it does not require Davidson to affirm the claim that an omniscient interpreter exists. According to the second formulation, Davidson is only committed to the claim that an omniscient interpreter could exist compatibly with our beliefs being what they are. Surely, this isn't something that strikes us as at all implausible. Indeed, many people do believe that an omniscient interpreter exists, and though it may be that they are not well justified in so believing, we can hardly accuse them of logical incoherence.

On the face of it, then, we have arrived at a plausible reconstruction of Davidson's Omniscient Interpreter Argument. In the next section of this chapter, I aim to evaluate this argument in light of Davidson's externalism. By digging a bit deeper, I think we shall be able to uncover some difficulties with premise [3].

VI: Davidson's Externalism

In "Knowledge of our Own Minds," Davidson offers the following thought experiment which is intended to bolster his externalism.

> Suppose lightning strikes a tree in a swamp; suppose that I am standing
> nearby. My body is reduced to its elements, while entirely by coincidence
> (and out of different molecules) the tree is turned into my physical replica.
> My replica, the Swampman, moves exactly as I did; according to its nature
> it departs the swamp, encounters and seems to recognize my friends, and
> appears to return their greetings in English. It moves into my house and
> seems to write articles on radical interpretation. No one can tell the
> difference.[12]

Davidson's contention, here, is that because the Swampman does not have the appropriate

causal history, it can neither understand English, nor can it have any thoughts at all.

Let us now consider a slight modification of Davidson's thought experiment.

Suppose lightning strikes everyone except yourself. Unbeknownst to you, everyone else's

body is reduced to its elements and, lo and behold, replicas turn up everywhere. With the

exception of yourself, there are no remaining language users on the face of the Earth, though,

of course, these facts escape your notice. We can suppose these replicas seem to go about

their business, according to their natures: swamp-students walk into your classrooms, sit

down, and appear to be taking notes, asking questions and the like; swamp-friends seem to

leave messages on your answering machine, and swamp-advisors only seem to demand the

next chapter. . . .

If all this were the case, would you not go on attributing beliefs to your swamp-

students, your swamp-friends, your swamp-colleagues, and to the other swamp-people you

encounter? Would you not go on interpreting them? Would you not find yourself supposing

that you are largely in agreement with them? Would you not go on supposing that the

---

[12] Davidson, "Knowing One's Own Mind," p. 443.

omniscient interpreter could interpret them? If so, you would be wrong. Swamp-people do not satisfy the conditions which, on Davidson's account, are necessary for speaking a language and possessing thoughts; therefore, they cannot be correctly interpreted. The omniscient interpreter would know this; however, you would not.

Consider the following argument.

[1] The OI is constrained by the same principles as a fallible interpreter. (premise)

[2] The fallible interpreter's job is to maximize the intelligibility of the speaker's utterances. (premise)

[3] The fallible interpreter does this by maximizing agreement between herself and the person she is interpreting. (premise)

It is clear from what has been said thus far that Davidson is committed to [1] - [3]. However, given these claims, it follows that:

[4] The OI can interpret me only if he maximizes agreement between himself and me--that is, only if he takes me to have beliefs which are similar to his own. (from 1, 2, and 3)

Now, suppose that the actual world is a swamp-world. If this were the case, then:

[5] The OI would know that those around me do not have the appropriate causal history which Davidson claims is necessary for possessing a language.

This follows from Davidson's definition of an Omniscient Interpreter and from his externalism. However, from [5] it follows that:

[6]   If the actual world were a swamp-world, the OI would know that those around me cannot be correctly interpreted. (from 5)

[7]   If the actual world were a swamp-world, then the OI would interpret me so as to put me largely in agreement with him. (from 4)

[8]   Therefore, he would interpret me as believing that those around me do not speak a language. (from 6 and 7)

Yet, it is clear that:

[9]   If the actual world were a swamp-world, I would go on falsely supposing that the people around me do speak a language. (premise)

[10]   Therefore, if the actual world were a swamp-world, then the OI would attribute beliefs to me which I do not possess--i.e., his assumption that my beliefs are similar to his own would be false.

The problem, then, is that if the actual world were a swamp-world, then the OI wouldn't get my beliefs right. If he could interpret me at all, he would do so incorrectly.[13]

---

[13] It may sound odd to say that an omniscient interpreter could get my beliefs wrong. In clarifying the notion of an omniscient interpreter, Vrinda Dalmiya remarks:

> The omniscient interpreter does not have privileged access into [the] minds of others and to know what we mean by our utterances he must go through the same process of interpretation as ordinary interpreters. Says Davidson, 'He [the omniscient interpreter] attributes beliefs to others, and interprets their speech on the basis of his own beliefs, just as the rest of us do'. . . . It seems to me that the crucial idea here is that of infallibility rather than that of 'omniscience.' The latter is 'all knowing' in the sense of knowing all truths. Davidson's interpreter, however, has to discover truth just like any ordinary being: thus, prior to the discovery he does not know all that there is to know. Maybe he will achieve omniscience as he acquires more and more beliefs but for a start his knowledge is characterized not in terms of all truths but in terms of truths only. *An omniscient interpreter then, cannot interpret incorrectly.* If he does interpret us and claims to know what we mean, he must be correct. [emphasis added]

(Vrinda Dalmiya, "Coherence, Truth and the 'Omniscient Interpreter'," pp. 87-88.) I quote the

Is this a good argument? The obvious point of contention is premise [8]. Why

couldn't Davidson simply deny [8]? Why couldn't Davidson simply bite the bullet and claim

---

paragraph in full because the term "then" in the italicized portion suggests that this is the conclusion of an argument.

Now, although I think that the central point in this paragraph is correct--that Davidson's interpreter is not, strictly speaking, omniscient--I do not think her claim that an OI can't get our beliefs wrong is true. She is correct in noting that

Davidson's interpreter is not omniscient. The OI does not have knowledge of everything. If he did have knowledge about everything, then he would not have to use the same method as fallible interpreters in determining the meaning of our words and the contents of our thoughts. He would already know them.

But, is Davidson's interpreter infallible, as Dalmiya suggests? Does he get our beliefs right by *fiat*? I don't think this supposition makes sense. Even the omniscient interpreter must make revisions in formulating a theory of interpretation for us. Would it thereby follow that our beliefs changed every time the OI found a need to revise his theory? Indeed, we can't even talk intelligibly about the OI's having a "completed" theory of interpretation, for it is clear that any number of future observations would force the OI to revise his theory. Insofar as the OI has to figure out what our beliefs are by observing the objects and events which he takes to prompt our assent to sentences, he is fallible. And, as Davidson claims, he has no choice but to infer what our words mean and what our beliefs are on the basis of such observations.

Moreover, in as much as he tries to make sense of our interaction with the external world, then it follows that the more divergence there is between our beliefs and the way the word is, the more he will attribute beliefs to us that we don't really possess. In short, the more we are mistaken, the more he will get our beliefs wrong.

There is something here that needs to be examined. Let's go back to Davidson's statement of the OI argument. "The omniscient interpreter, using the same method as the fallible interpreter, finds the fallible speaker largely consistent and correct. By his own standards, of course, but since these are objectively correct, the fallible speaker *is seen to be* largely consistent and correct by objective standards" [emphasis added]. Why does Davidson use the more round-about "is seen to be" instead of just saying that the fallible speaker *is* largely consistent and correct? It is in answering this question that we begin to see a fundamental problem with Davidson's response to the skeptic.

There is a tension here. To make sense of Davidson's argument, we need to view the OI as using the same method of interpretation that we use. We can make sense of interpreting another only if we know that the beliefs we attribute to others are, in part, our constructs. Our theory of interpretation may or may not be correct. We need to always be willing to concede that we have interpreted incorrectly. We don't get to look into another person's head and read off her beliefs. The whole notion of maximizing the intelligibility of the utterances of others makes sense only in the face our recognizing that we can't just read off the beliefs of others from what we take to be the causes of their beliefs. Now, if we are to make sense of the OI's using the same method we use, then we must assume that the OI could get our beliefs wrong.

that the OI could still interpret us (correctly) even if the actual world were a swamp-world? To run this line, Davidson would have to say that an OI could interpret a speaker even though that speaker has false beliefs in an entire domain--in this case, every belief the speaker has about other minds is false.

I don't think this is a plausible line of defense. For one thing, if Davidson were to take this line, he would have to revise or at least modify his claim that "massive error" is unintelligible. After all, each of us has a good many beliefs about those in our lives. For example, I believe that the members of my dissertation committee would have desired to have this dissertation a bit sooner. I am fully convinced that my parents will feel quite happy when I finally get a job. But if neither the members of my dissertation committee, nor my parents have minds with which to do any thinking or believing, then a good many of my beliefs are false. In short, I would be massively mistaken about the nature of my environment.

Although I do not think this objection is conclusive, I do think it highlights an important question, one which Davidson has yet to answer--namely, the question "how wrong must we be about our environment in order to deter interpretation?" In pursuing this question, we shall be led to a deeper criticism of Davidson's Omniscient Interpreter Argument.

V: Are there Empirical Presuppositions in Davidson's OI Argument?

Let us now return to Davidson's Omniscient Interpreter Argument. Recall that we have been interested in the question of whether premise [3], the claim that our beliefs are

such that if there were an OI, he could interpret us, tacitly relies on any empirical presuppositions. In the previous section, I posed the following question: if the actual world were a swamp-world, could the OI still interpret us? In other words, if the actual world were a swamp-world would our beliefs be such that an OI could interpret us, if one were to exist?

Posing the question in this way reveals that premise [3] of Davidson's OI Argument is ambiguous. As stated premise [3] reads, "our beliefs are such that if there were an OI, he could interpret us." This admits of two different readings.

(1) Our beliefs are such that if there were an omniscient interpreter, then there is some possible world in which he could interpret us.

(2) Our beliefs are such that if there were an omniscient interpreter, then he could interpret us *as we actually are*.

If we read premise [3] in the sense of (1), then it is trivially true: if the actual world were a swamp-world, then there would still exist a possible world in which our beliefs turn out mostly true. But, surely, Davidson doesn't intend [3] to be read in the sense of (1). It is clear that for this argument to be even so much as relevant to the problem of skepticism, we have to read premise [3] along the lines of (2) and not (1). Indeed, Davidson's thought which motivated this line of reasoning was that given that we are language speakers, and that to be language speakers we have to be interpretable *as we actually are*, then if there were an omniscient interpreter, he would be able to interpret us.

But what does this "actually are" actually come to? To get the right reading of [3], Davidson needs to claim that our beliefs and situation are such that if there were an OI, he

could interpret us--that is, that he could interpret us in our actual environment. To take this into consideration, we need to revise the OI Argument along the following lines.

[1] Possibly, there exists an omniscient interpreter--a being whose beliefs about the world are all true. (premise)

[2] For any x and for any y, x can interpret y (correctly) only if x's beliefs are largely in agreement with y's. (premise)

[3] Our beliefs and situation are such that if there were an omniscient interpreter, he could interpret us. (premise)

[4] Therefore, our beliefs are largely in agreement with the beliefs that an omniscient interpreter would have, were one to exist. (from 2 and 3)

[5] Therefore, our beliefs are largely true. (from 1 and 4)

But notice that to affirm [3], it looks as though we need to know something about our actual environment. To see this, recall that it is part of Davidson's method of interpretation that the OI (or any interpreter, for that matter) interpret the speech of another on the basis of public cues. As Davidson sets out the OI Argument, the OI is not so omniscient that he already knows what our beliefs are and what our words mean. Like every other interpreter, he must take the objects of our beliefs to be the causes of our beliefs. Recall a passage quoted earlier:

> if I am right, we can't in general first identify beliefs and meanings and then ask what caused them. The causality plays an indispensable role in determining the content of what we say and believe.[14]

---

[14] Davidson, "A Coherence Theory of Truth and Knowledge," p. 317.

Given this, it is apparent that we can affirm [3]--that our beliefs are such that an OI could interpret us, were one to exist--only if we already know that we live in a world which admits of causal relations. But this is an empirical claim. Hence, it is not one to which Davidson can legitimately appeal in arguing against the skeptic.

Thus, like Burge and Putnam, the most Davidson can draw is a conclusion of conditional form: if we live in a world which admits of causal relations, then our thoughts and the meanings of our terms are about what causes them. As Davidson himself has stated the conditional: "If words and thoughts are, in the most basic cases, necessarily about the sorts of objects and events that cause them, there is no room for Cartesian doubts about the independent existences of such objects and events."[15] As we have observed in our discussion of Burge and Putnam, however, to affirm the antecedent of this conditional is to beg the question against the skeptic.

VI: Davidson's Argument from Objectivity

In this section, we turn to some later developments in Davidson's thought. We shall be concerned with the anti-skeptical argument he presents in "The Second Person" and "Three Varieties of Knowledge." In these papers, he develops an argument from the necessary public character of language to the conclusion that skepticism about the external world and skepticism about other minds are false.

Davidson begins by asking the question: "what are the conditions under which one can be said to possess a belief?" His answer is this:

---

[15] Davidson, "The Myth of the Subjective," p. 165.

[T]o have a belief it is not enough to discriminate among aspects of the world, to behave in different ways in different circumstances; an earthworm or a sunflower does this. Having a belief demands in addition appreciating the contrast between true belief and false, between appearance and reality, mere seeming and being. We can, of course, say that a sunflower has made a mistake if it turns toward an artificial light as if it were the sun, but we do not suppose the sunflower can think it has made a mistake, and so we do not attribute a belief to the sunflower. Someone who has a belief about the world--or anything else--must grasp the concept of objective truth, of what is the case independently of what he or she thinks.[16]

Davidson's claim, then, is that to possess a belief one must possess the concept of objective truth.

Next, Davidson asks the question: "How do we come to possess the concept of objective truth?" His answer is that we come to possess the concept of objective truth through interpersonal communication. "The source of the concept of objective truth is interpersonal communication. Thought depends on communication. This follows at once if we suppose that language is essential to thought, and we agree with Wittgenstein that there cannot be a private language."[17] On Davidson's reading of Wittgenstein, the reason that there cannot be a private language is that there would be no way to distinguish correct from incorrect uses of words. "If only communication can provide a check on the correct use of words, only communication can supply a standard of objectivity."[18]

There are two claims that need to be distinguished here. The first is that the source of objective truth is interpersonal communication. The second is that the *only* source of

---

[16]

[17] Davidson, "Three Varieties of Knowledge," p. 157.

[18] *Ibid.*, p. 157.

objective truth is interpersonal communication. I want to draw attention to this distinction

because I think that Davidson actually needs the second claim, which is much stronger than

the first. But on the face of it, the stronger claim is false. Indeed, it looks as though we can

acquire the concept of objective truth simply by correcting errors in our own belief system.

Nonetheless, let us waive this objection for now.

Davidson's next claim is that interpersonal communication is possible only if

speakers and interpreters share beliefs about the world which are largely true. Why is this?

Davidson's claim seems to be that two speakers can communicate only if they are reacting

to the same stimuli in their environment. Davidson puts the point as follows:

> Without this sharing of reactions to common stimuli, thought and speech
> would have no particular content--that is, no content at all. It takes two
> points of view to give a location to the cause of a thought, and thus to define
> its content. We may think of it as a form of triangulation: each of two people
> is reacting differentially to sensory stimuli streaming in from a certain
> direction. If we project the incoming lines outward, their intersection is the
> common cause. If thee two people now notes each other's reactions . . . each
> can correlate these observed reactions with his or her stimuli from the world.
> The common cause can now determine the contents of an utterance and a
> thought. The triangle which gives content of thought and speech is complete.
> But it takes two to triangulate. Two, or, of course, more.[19]

Davidson takes triangulation to be a necessary condition on having any thoughts at all. For

without convergence on a common cause, there would be no way to identify a particular

object as the object of thought. As Davidson has put the point, "Without such sharing, there

---

[19] *Ibid.*, pp. 159-160.

would be no grounds for selecting one cause rather than another as the content-fixing cause."[20]

Let me pause to sum up the state of the argument. Thus far, Davidson has argued for three claims: (1) that one can possess a belief only if one possesses the concept of objective truth, (2) that one can acquire the concept of objective truth only through interpersonal communication, and (3) interpersonal communication is possible only if two speakers share beliefs about the external world which are mostly true. Given these three claims, it is a small step to establishing the conclusion that skepticism about the external world and skepticism about other minds is false. The argument may be stated as follows.

[1]   It is possible for one to possess a belief only if one possesses the concept of objective truth.

[2]   The concept of objective truth can be acquired only through interpersonal communication.

[3]   Two speakers can engage in interpersonal communication only if they share beliefs about their environment which are mostly true.

[4]   I have beliefs.

[5]   Therefore, I possess the concept of objective truth. (from 1 and 4)

[6]   Therefore, I do engage in interpersonal communication. (from 2 and 5)

[7]   Therefore, I share beliefs with others about my environment that are mostly true. (from 3 and 6)

[8]   If [7], then skepticism about other minds and skepticism about the external world is false.

---

[20] Davidson, "Epistemology Externalized," p. 194.

[9]   Hence, skepticism about other minds and skepticism about the external world is false. (from 7 and 8)

Is this a good argument? In the next section, we shall consider three objections to this argument.

VII. Three Objections to Davidson's Argument from Objectivity

There are good reasons to think that premises [1], [2] and [3] are false. Let us begin with premise [1]. One reason to think premise [1] is false is that we routinely attribute beliefs to young children; yet, it is doubtful that the average five year old has the concept of objective truth. Indeed, the average five year old cannot discriminate between real events and fictional events. To say that one must possess the concept of objective truth in order to possess any beliefs at all is much too strong. Moreover, it is not uncommon for people to attribute beliefs and desires to their pets. For example, I am fully convinced that my cat desires cat treats, that she believes that when I shake the can of Pounce, that she believes she is going to get a cat treat and so forth. I am quite willing to say that my cat has beliefs and desires; but as intelligent as I think she is, I do not think she possesses the concept of objective truth.

Premise [2] looks even more suspect. As noted above, to get a valid argument Davidson needs the claim that interpersonal communication is the *only* possible means of acquiring the concept of objective truth. Yet this is clearly false, since it is not unreasonable to suppose that a person could acquire the concept of objective truth simply by discovering and correcting inconsistencies in her own belief system. Moreover, there is nothing

incoherent about supposing that an evil genius placed the concept of objective truth in our minds. Suppose, also, that you were the only language user in a world filled with swamp people. If this were the case, you would never actually succeed in communicating with another person; yet, it seems reasonable to suppose that you could still acquire the concept of objective truth.

Premise [3] is also highly questionable. Why must people have largely true beliefs about the external world in order to communicate? On the face of it, agreement on general truths seems to be sufficient to allow for communication. After all, it seems plausible to say that two people trapped in virtual reality could communicate with each other. They could communicate with each other even though a significant portion of their beliefs about their environment were false.

VIII: Objectivity and Skepticism

In addition, we ought to examine the sorts of intuitions that Davidson seeks to elicit in introducing the notion of triangulation. Let us consider a passage from "The Second Person."

> . . . if I am right, the kind of triangulation I have described, while not sufficient to establish that a creature has a concept of a particular object or kind of object, is necessary if there is to be any answer at all to the question of what its concepts are concepts of. If we consider a single creature by itself, its responses, no matter how complex, cannot show that it is reacting to, or thinking about, events a certain distance away rather than, say, on its skin.[21]

---

[21] Davidson, "The Second Person," p. 263.

Triangulation can happen only in worlds that admit of causal relations. Indeed, part of the story about triangulation is that the interpreter and speaker react to each other's reaction to a common object or event. This seems to presuppose that we are physical creatures, living in a spatio-temporal world. It seems to presuppose that we are the sorts of creatures that are well adapted through evolution to that each of us react in similar fashion to similar stimuli and that each of us categorize objects and events in much the same way. Thus we can see that Davidson's talk of triangulation tacitly appeals to empirical presuppositions.

CHAPTER VI
CONCLUSION

Our aim throughout much of this dissertation has been to answer the following questions: (1) Is externalism true? and (2) Supposing that externalism is true, does it provide an adequate response to the skeptical challenge? I have argued that the answer to both of these questions is "no." This final chapter has two central aims. The first aim is to summarize the main arguments in support of this negative answer. The second, more ambitious aim is to assess the implications of this negative answer for resolving the skeptical challenge.

In chapter III, we took up two forms of externalism advanced by Tyler Burge: social externalism and perceptual externalism. We found that Burge's argument for social externalism falls prey to the traveling bilingual argument, which suggests that the principles of belief attribution sanctioned by social externalism lead to a contradiction: that we could have a single belief with two semantically distinct contents. We found Burge's argument for perceptual externalism is unconvincing, at least for those who do not share Burge's intuitions about the plausibility of individualism. On these grounds, we rejected social and perceptual externalism.

More interestingly, we found that even if these accounts of belief attribution were correct, they would be of no help in resolving the skeptical challenge. The most that we

could get from these theories is a conditional conclusion: if we live in a non-solipsistic, spatio-temporal world, then relational facts are relevant to the determination of thought content. However, as noted above, to argue from externalism to the falsity of skepticism about the external world, the externalist would have to affirm the antecedent of this conditional. Yet to do this is to presuppose that we already have knowledge that the skeptic denies.

In chapter IV, we took up the views of Hilary Putnam. The primary difficulty with Putnam's anti-skeptical argument lies in his appeal to the Causal Theory of Reference (CTR). Putnam, himself, has realized that the appeal to CTR in giving an anti-skeptical knowledge begs the question at hand. This is because we could know that CTR applies to us only if we first know that we live in a physically possible world. Yet to suppose that we know this is to assume that we already know certain facts about the external world.

In chapter V, we examined the views of Donald Davidson. We found that his primary anti-skeptical argument, the Omniscient Interpreter Argument, contains an equivocation. The difficulty focuses on his use of the term "interpretable" in the claim that we would be interpretable by an Omniscient Interpreter, were one to exist. This premise can be cashed out in two different ways: (1) there is some possible world in which we could be interpreted by an Omniscient Interpreter, were one to exist; and (2) the actual world is a possible world in which we could be interpreted by an Omniscient Interpreter, were one to exist. To get a valid argument, Davidson needs to cash out the interpretability assumption along the lines of (2).

Yet, when we examine the conditions under which a language user is interpretable (when we examine Davidson's externalism), we find that a language user can be interpreted only if it has the right sort of historical connections--the very sorts of connections which the Swampman lacked. We rejected Davidson's appeal to externalism in his anti-skeptical argument on the grounds that we could know that we are interpretable as we actually are only if we first knew that we had the right historical connections with our environment. For to know this is to presuppose that we already know certain facts about the external world.

This concludes the summary of the main arguments advanced in support my rejection of externalism. Let us, then, go on to the second aim of this final chapter--to assess the implications of this rejection of externalism. Why does externalism fail in general? It is one thing to pinpoint errors in particular anti-skeptical arguments; it is another thing to identify a general mistake that externalists make in using their theories against the skeptic.

In general, externalism is a theory about the conditions under which a language user can be said to have thoughts with certain specified contents. The problem is that giving an analysis of such conditions is not the same thing as giving an argument which establishes that these conditions obtain. At best, the externalist can arrive at a conclusion that says that relational factors are relevant to the determination of thought content, provided that we live in a world that admits of relational factors. But to say that we know that the actual world is a world that admits of relational facts is tantamount to saying that we already know that the actual world is a physically possible world.

It should be clear, then, that externalism could never establish the falsity of (M)--the assumption that there is no necessary connection between the contents of our thoughts and

the actual state of our environment. To establish this, the externalist would have to establish the claim that it is not possible for us to have the thoughts we do unless we lived in a physically possible world of a rather specific sort--i.e., one in which the postulated relational facts obtained. Yet no amount of conceptual analysis can yield this conclusion, for to apply the externalist conclusion against the skeptic would have to have antecedent knowledge of facts in the domain in question. If I am correct in my analysis of the assumptions underlying philosophical skepticism and in my analysis of the limited, conditional nature of the externalist conclusion, it follows there is no non-question begging response to the skeptical challenge to be found on the basis of externalism about thought content.

BIBLIOGRAPHY

Armstrong, D. (1961). *Perception and the Physical World*, London: Routledge and Kegan Paul.

_____. (1968). *A Materialist Theory of Mind*, New York: Humanities Press.

_____. (1973). *Belief, Truth and Knowledge*, Cambridge: Cambridge University Press.

_____. (1981). *The Nature of Mind and Other Essays*, New York: Cornell University Press.

Ayer, A.J. (1946). *Language, Truth and Logic*, London: Gollancz.

Bach, K. (1982). "*De Re* Belief and Methodological Solipsism," in A. Woodfield (ed.), *Thought and Object*, Oxford: Clarendon Press, 121-153.

_____. (1987). *Thought and Reference*, Oxford: Clarendon Press.

_____. (1988). "Burge's New Thought Experiment: Back to the Drawing Room," *Journal of Philosophy*, 85: 88-97.

Baker, L. (1987). "Content by Courtesy," *Journal of Philosophy*, 84: 197-213.

_____. (1989). "On a Causal Theory of Content," *Philosophical Perspectives*, 3: 165-186.

_____. (1991). "Dretske on the Explanatory Role of Belief," *Philosophical Studies*, 63: 99-111.

Berkeley, G. (1710). *A Treatise Concerning the Principles of Human Knowledge* in M.R. Ayers, (ed.), *Berkeley: Philosophical Work*, London: Dent, 1975.

_____. (1713). *Three Dialogues Between Hylas and Philonous* in M.R. Ayers, (ed.), *Berkeley: Philosophical Work*, London: Dent, 1975.

Brueckner, A. (1986a). "Brains in a Vat," *Journal of Philosophy*, 83: 148-167.

_____. (1998b). "Charity and Skepticism," *Pacific Philosophical Quarterly*, 67: 264-268.

_____. (1990). "Skepticism about Knowledge of Content," *Mind*, 99: 447-451.

_____. (1991). "The Omniscient Interpreter Rides Again," *Analysis*, 51: 199-205.

_____. (1992a). "If I am a Brain in a Vat, then I am not a Brain in a Vat," *Mind*, 101: 123-128.

_____. (1992b). "What an Anti-Individualist Knows A Priori," *Analysis*, 52: 111-118.

_____. (1992c). "Conceiving One's Envatment while Denying Metaphysical Realism," *Australasian Journal of Philosophy*, 70: 469-474.

_____. (1993). "Skepticism and Externalism," *Philosophia*, 22: 169-171.

_____. (1994). "Knowledge of Content and Knowledge of the World," *Philosophical Review*, 103: 327-343.

_____. (1995a). "The Characteristic Thesis of Anti-Individualism," *Analysis*, 55: 146-148.

_____. (1995b). "Skepticism and the Causal Theory of Reference," *Philosophical Quarterly*, 45: 199-201.

Buekens, F. (1994). "Externalism, Content, and Causal Histories," *Dialectica*, 48: 267-286.

Burge, T. (1977). "Belief *De Re*," *Journal of Philosophy*, 74: 338-363.

_____. (1979). "Individualism and the Mental," in *Midwest Studies in Philosophy, vol. IV: Studies in Metaphysics*, P.A. French, T.E. Uehling and H.K. Wettstein (eds.), Minneapolis: University of Minnesota Press, 73-121. Reprinted in D.M. Rosenthal (ed.), *The Nature of Mind*, Oxford: Oxford University Press, 1991, pp. 536-589.

_____. (1982). "Two Thought Experiments Reviewed," *Notre Dame Journal of Formal Logic*, 23: 284-293.

_____. (1982). "Other Bodies," in A. Woodfield (ed.), *Thought and Object: Essays on Intentionality*, Oxford: Oxford University Press, pp. 97-120.

_____. (1986). "Cartesian Error and the Objectivity of Perception," in P. Pettit and J. McDowell (eds.), *Subject, Thought, and Context*, Oxford: Oxford University Press, 1986, pp. 117-136.

_____. (1986). "Intellectual Norms and the Foundations of Mind," *Journal of Philosophy*, 83: 697-720.

_____. (1986). "Individualism and Psychology," *Philosophical Review*, 95: 3-46.

_____. (1988). "Individualism and Self-Knowledge," *Journal of Philosophy*, 85: 649-663.

_____. (1993). "Mind-Body Causation and Explanatory Practice," in J. Heil and A. Mele (eds.), *Mental Causation*, Oxford: Clarendon Press, 1993, pp. 96-120.

_____. (1993). "Content Preservation," *Philosophical Review*, 102: 457-488.

Casati, R. and J. Dokic. (1991). "Brains in a Vat, Language and Meta-Language," *Analysis*, 51: 91-93.

Collier, J. (1990). "Could I Conceive Being a Brain in a Vat?" *Australasian Journal of Philosophy*, 68: 413-419.

Carrier, L.S. (1981). *Experience and the Objects of Perception*, Washington: University Press of America.

Cottingham, J. (1992). "Descartes," in J. Dancy and E. Sosa (eds.), *A Companion to Epistemology* (Oxford: Blackwell), pp. 93-97.

Craig, E. (1990). "Davidson and the Skeptic: The Thumbnail Version," *Analysis*, 50: 213-218.

David, M. (1991). "Neither Mentioning `Brains in a Vat' nor Mentioning Brains in a Vat Will Prove that We are not Brains in a Vat," *Philosophy and Phenomenological Research*, 101: 891-896.

Davidson, D. (1973). "Radical Interpretation," *Dialectica*, 27: 313-328.

_____. (1973). "Belief as the Basis of Meaning," *Synthese*, 27: 309-323.

_____. (1974). "On the Very Idea of a Conceptual Scheme," *Proceedings and Addresses of the American Philosophical Association*, 47: 5-25.

_____. (1977). "The Method of Truth in Metaphysics" in P.A. French, J.R. Uehling, and H.K. Wettstein (eds.), *Midwest Studies in Philosophy*, 2: 244-254.

_____. (1980). *Essays on Actions and Events*, Oxford: Clarendon Press.

_____. (1982). "Rational Animals," *Dialectica*, 36: 317-327.

_____. (1982). "Empirical Content" in E. Lepore (ed.), *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, Blackewll, Oxford, 1986.

_____. (1983). "A Coherence Theory of Truth and Knowledge" in D. Henrich (ed.), *Kant oder Hegel*, Stuttgart, Klett-Cotta. Reprinted in E. Lepore (ed.), *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, Blackewll, Oxford, 1986.

_____. (1984). *Inquiries into Truth and Interpretation*, Oxford: Clarendon Press.

_____. (1984). "Thought and Talk," in *Inquiries into Truth and Interpretation*, Oxford: Clarendon Press.

_____. (1984). "First-Person Authority," *Dialectica*, 38: 101-111.

_____. (1987). "Knowing One's Own Mind," *Proceedings of the American Philosophical Association*, 60: 441-458.

_____. (1988). "Reply to Burge," *Journal of Philosophy*, 85: 664-665.

_____. (1989). "Conditions of Thought" in J. Brandl and W. Gombocz (eds.), *The Mind of Donald Davidson*, Amsterdam: Rodopi.

_____. (1990). "Representation and Interpretation" in W.H. Newton-Smith and K.V. Wilkes (eds.), *Modelling the Mind*, Oxford: Oxford University Press.

_____. (1991). "Epistemology Externalized," *Dialectica*, 45: 191-202.

_____. (1991). "Three Varieties of Knowledge" in A. Griffiths (ed.), *A.J. Ayer: Memorial Essays*, Cambridge: Cambridge University Press.

_____. (1992). "The Second Person," *Midwest Studies in Philosophy*, XVII: 255-267.

_____. (1993). "Thinking Causes," in J. Heil and A. Mele (eds.), *Mental Causation*, Oxford: Clarendon Press, pp. 3-17.

_____. (1993). "Interpreting Radical Interpretation," in J. Tomberlin, (ed.), *Philosophical Perspectives*, Atascadero: Ridgeview.

Davies, M. (1991). "Individualism and Perceptual Content," *Mind*, 100: 461-484.

_____. (1992). "Perceptual Content and Local Supervenience," *Proceedings of the Aristotelian Society*, 92: 21-45.

_____. (1993). "Aims and Claims of Externalist Arguments," *Philosophical Issues*, 4: 227-249.

Dennett, D. (1982). "Beyond Belief," in A. Woodfield (ed.), *Thought and Object*, Oxford: Clarendon Press, 1-98.

Descartes, R. (1641). *Meditations on First Philosophy* in *The Philosophical Works of Descartes*, in E. Haldane and G. Ross, (trans.), Cambridge: Cambridge University Press, 1931.

Donnellan, K. (1966). "Reference and Definite Descriptions," *Philosophical Review*, 75: 281-304.

Dretske, F. (1981). *Knowledge and the Flow of Information*, Cambridge: MIT Press.

_____. (1986). "Misrepresentation," in R. Bogdan (ed.), *Belief: Form, Content, and Function*, New York: Oxford, 17-36.

_____. (1993). "Mental Events as Structuring Causes of Behavior," in J. Heil and A. Mele (eds.), *Mental Causation*, Oxford: Clarendon Press, 121-136.

_____. (1995). *Naturalizing the Mind*, Cambridge: MIT Press.

Dummett, M. (1973). *Frege: Philosophy of Language*, New York: Harper and Row.

_____. (1978). *Truth and Other Enigmas*, Cambridge: Harvard University Press.

Evans, G. (1982). *The Varieties of Reference*, Oxford: Clarendon Press.

Falvey, K. and J. Owens. (1994). "Externalism, Self-Knowledge, and Skepticism," *Philosophical Review*, 103: 107-137.

Farrell, F. (1986). "Putnam and the Vat-People," *Philosophia*, 16, 147-160.

Feldman, S. (1984). "Refutation of Dogmatism: Putnam's Brains in Vats," *Southern Journal of Philosophy*, 22: 323-330.

Foley, R. (1990). "Skepticism and Rationality," in M.R. Roth and G. Ross (eds.), *Doubting: Contemporary Perspectives on Skepticism*, Dordrecht: Kluwer, 69-82.

Foley, R. and R. Fumerton. (1985). "Davidson's Theism?" *Philosophical Studies*, 48: 83-89.

Forbes, G. (1995). "Realism and Skepticism: Brains in a Vat Revisited," *Journal of Philosophy*, 92: 205-222.

Frankfurt, H. (1970). *Demons, Dreamers and Madmen: The Defense of Reason in Descartes' Meditations,* New York: Bobbs-Merrill.

Frege, G. (1892). "On Sense and Reference," in P.T. Geach and M. Black (eds.), *Translations for the Philosophical Writings of Gottlob Frege*, Oxford: Blackwell.

Fumerton, R. (1990). "Metaepistemology and Skepticism," in M.R. Roth and G. Ross (eds.), *Doubting: Contemporary Perspectives on Skepticism*, Dordrecht: Kluwer, 57-68.

Gallois, A.N. (1992). "Putnam, Brains in Vats and Arguments for Skepticism," *Mind*, 101: 273-285.

Georgalis, N. (1990). "No Access for the Externalist: Discussion of Heil's `Privileged Access'," *Mind*, 99: 101-108.

Genova, A.C. (1991). "Craig on Davidson: A Thumbnail Refutation," *Analysis*, 50: 195-198.

Getteir, E. (1963). "Is Justified True Belief Knowledge?" *Analysis*, 23: 121-123.

Grayling, A.C. (1986). *Berkeley: The Central Arguments*, London: Routledge.

Goldman, A.I. (1979). "What is Justified Belief?" in Pappas (ed.), *Justification and Knowledge*, Dordrecht: Reidel, pp. 1-23.

Guttenplan, S. (1994). "Belief, Knowledge and the Orgins of Content," *Dialectica*, 48: 287-305.

Harrison, J. (1985). "Professor Putnam on Brains in Vats," *Erkenntnis*, 23: 55-57.

Heil, J. (1988). "Privileged Access," *Mind*, 97: 238-251.

_____. (1992). *The Nature of True Minds*, Cambridge: Cambridge University Press.

_____ and A. Mele (eds.). (1993). *Mental Causation*, Oxford: Clarendon Press.

Hume, D. (1739). *A Treatise of Human Nature*, L.A. Selby-Bigge, (ed.), Oxford: Oxford University Press, 1978.

_____. (1748). *An Enquiry Concerning Human Understanding*, L.A. Selby-Bigge, (ed.), Oxford: Oxford University Press, 1975.

Iseminger, G. (1988). "Putnam's Miraculous Argument," *Analysis*, 48: 190-195.

Kant, I. (1787). *Critique of Pure Reason*, N. Kemp Smith, (trans.), London: Macmillan, 1964).

Kinghan, M. (1986). "The External World Skeptic Escapes Again," *Philosophia*, 16: 161-166.

Kirk, R. (1994). "The Trouble with Ultra-Externalism," *Proceedings of the Aristotelian Society*, 94: 293-307.

Klein, P. (1981). "Radical Interpretation and Global Skepticism" in *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, in E. LePore, (ed.), New York: Blackwell, 1986.

Kripke, S. (1977). "Speaker's Reference and Semantic Reference," in P. A. French, T. E. Uehling and H. K. Wettstein (eds.), *Contemporary Perspectives in the Philosophy of Language*, Minneapolis: University of Minnesota Press, 6-27.

_____. (1979). "A Puzzle about Belief," in A. Margalit (ed.), *Meaning and Use*, Dordrecht: Reidel, 239-283.

LePore, E. (ed.). (1986). *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, New York: Blackwell.

_____ and B. McLaughlin (eds.). *Actions and Events: Perspectives on the Philosophy of Donald Davidson*, Oxford: Blackwell.

Lewis, D. (1981). "What Puzzling Pierre does not Believe," *Australasian Journal of Philosophy*, 59: 283-289.

Ludlow, P. (1995). "Externalism, Self-Knowledge, and the Prevalence of Slow Switching," *Analysis*, 55: 45-49.

_____, (1995). "Social Externalism, Self-Knowledge, and Memory," *Analysis*, 55: 157-159.

Ludwig, K. (1992). "Brains in a Vat, Subjectivity, and the Causal Theory of Reference," *Journal of Philosophical Research*, 17: 314-345.

_____. (1992). "Singular Thoughts and the Cartesian Theory of Mind," *Nous*, 30: 434-456.

_____. (1992). "Skepticism and Interpretation," *Philosophy and Phenomenological Research*.

_____. (1993). "Externalism, Naturalism and Method," *Philosophical Issues*, 4: 250-264.

_____. (1994). "Causal Relevance and Thought Content," *Philosophical Quarterly*, 44: 334-353.

_____. (1996). "Duplicating Thoughts," *Mind and Language*, 11: 92-102.

Luper-Foy, S. (1988). "The Knower, Inside and Out," *Synthese*, 47: 349-367.

MacDonald, C. (ed.). (1995). *Philosophy of Psychology: Debates on Psychological Explanation*, Cambridge, Blackwell.

_____ and G. MacDonald. (1986). "Mental Causes and Explanation of Action,"*Philosophical Quarterly*, 36: 145-158.

Maker, W. (1991). "Davidson's Transcendental Arguments," *Philosophy and Phenomenological Research*, 51: 345-360.

Martin, M.G.F. (1992). "Perception, Concepts and Memory," *Philosophical Review*, 101: 745-763.

McCulloch, G. (1993). "The Very Idea of the Phenomenological," *Proceedings of the Aristotelian Society*, 93: 39-57.

McDowell, J. (1994). "The Content of Perceptual Experience," *Philosophical Quarterly*, 44: 190-205.

_____. (1994). "Not Much Trouble for Ultra-Externalism," *Analysis*, 54: 265-269.

McGinn, C. (1982). "The Structure of Content," in A. Woodfield (ed.), *Thought and Object*, Oxford: Clarendon Press, 207-258.

_____. (1986). "Radical Interpretation and Epistemology," in *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, E. LePore, ed., (New York: Blackwell, 1986).

_____. (1991). *Mental Content*, Cambridge: Blackwell.

McIntyre, J. (1984). "Putnam's Brains," *Analysis*, 44: 59-61.

Millar, A. (1986). "What's in a Look," *Proceedings of the Aristotelian Society*, 86: 83-97.

_____. (1991). "Concepts, Experience and Inference," *Mind*, 100: 495-504.

Moran, R. (1994). "Interpretation Theory and the First Person," *Philosophical Quarterly*, 44: 154-173.

Noonan, H. (1991). "Object-Dependent Thoughts and Psychological Redundancy," *Analysis*, 51: 1-16.

Nozick, R. (1981). *Philosophical Explanations*, Cambridge: Harvard University Press.

Owens, J. (1994). "Psychological Externalism and Psychological Explanation," *Philosophy and Phenomenological Research*, 54: 921-928.

_____. (1987). "In Defense of a Different Doppelganger," *Philosophical Review*, 96: 521-554.

Peacocke, C. (1983). *Sense and Content*, Oxford: Oxford University Press.

_____. (1986). *Thoughts: An Essay on Content*, New York: Blackwell.

_____. (1992). *A Study of Concepts*, Cambridge: MIT Press.

_____. (1993). "Externalism and Explanation," *Proceedings of the Aristotelian Society*, 93: 203-230.

Pessin, A. (ed.). (1995). *The Twin-Earth Chronicles*, London: Paragon House.

Pettit, P. and J. McDowell (eds.). (1986). *Subject, Thought, and Context*, Oxford: Clarendon Press.

Putnam, H. (1975). "The Meaning of `Meaning,'" in *Mind Language and Reality: Philosophical Papers, Vol. II*, Cambridge: Cambridge University Press.

_____. (1981). *Reason, Truth and History*, Cambridge: Cambridge University Press.

_____. (1992). "Is it Necessary that Water is H₂0?" in *The Philosophy of A.J. Ayer*, L.E. Hahn, ed. (La Salle: Open Court).

Preti, C. (1992). "Opacity, Belief and Analyticity," *Philosophical Studies*, 66: 297-306.

_____. (1995). "Externalism and Analyticity," *Philosophical Studies*, 79: 213-236.

Popkin, R. (1979). *The History of Skepticis from Erasmus to Descartes* (Berkeley: University of California Press).

Quine, W.V.O. (1951). "Two Dogmas of Empiricism," *Philosophical Review*. Reprinted in *From A Logical Point of View*, Cambridge, Harvard University Press, pp. 20-46.

_____. (1953). *From A Logical Point of View*, Cambridge, Harvard University Press.

_____. (1960). *Word and Object*, Cambridge: MIT Press.

_____. (1969). *Ontological Relativity and Other Essays*, New York: Columbia University Press.

_____. (1990). *The Persuit of Truth*, Cambridge: Harvard University Press.

Rey, G. (1992). "Semantic Externalism and Conceptual Competence," *Proceedings of the Aristotelian Society*, 92: 315-333.

Robinson, H. (1992). "Experience and Externalism: A Reply to Peter Smith," *Proceedings of the Aristotelian Society*, 92: 221-223.

Sartwell, C. (1995). "Radical Externalism Concerning Experience," *Philosophical Studies*, 78: 55-70.

Schiffer, S. (1992). "Boghossian on Externalism and Inference," *Philosophical Issues*, 2: 29-37.

Segal, G. (1989). "Seeing What is Not There," *Philosophical Review*, 98: 189-214.

_____. (1989). "The Return of the Individual," *Mind*, 100: 1-17.

_____. (1991). "Defence of a Reasonable Individualism," *Mind*, 100: 485-494.

_____ and S. Elliott. (1991). "The Causal Efficacy of Content," *Philosophical Studies*, 63: 1-30.

Seymour, D. (1993). "Some of the Difference in the World: Crane on Intentional Content," *Philosophical Quarterly*, 43: 83-89.

Smith, P. (1984). "Could We be Brains in a Vat?" *Canadian Journal of Philosophy*, 14: 115-123.

Sosa, E. (1993). "Abilities, Concepts, and Externalism," in J. Heil and A. Mele (eds.), *Mental Causation*, Oxford: Clarendon Press, 309-328.

Stamp, D.W. (1979). "Toward a Causal Theory of Linguistic Representation," in P. French, T. Uehling and H. Wettstein (eds.), *Contemporary Perspectives in the Philosophy of Language*, Minneapolis: University of Minnesota Press, 81-102.

_____. (1986). "Verificationism and a Causal Theory of Meaning," *Synthese*, 69: 107-137.

Steinitz, Y. (1994). "Brains in a Vat: Different Perspectives," *Philosophical Quarterly*, 44: 213-222.

Stephens, J. and L. Russow. (1985). "Brains in Vats and the Internalist Perspective," *Australasian Journal of Philosophy*, 63: 205-212.

Stich, S.P. (1982). "On the Ascription of Content," in A. Woodfield (ed.), *Thought and Object: Essays on Intentionality*, Oxford: Oxford University Press, 153-206.

Stroud, B. (1968). "Transcendental Arguments," *Journal of Philosophy*, 65: 241-256.

_____. (1981). "The Significance of Naturalized Epistemology," in H. Kornblith, (ed.), *Naturalizing Epistemology* Cambridge: MIT Press.

_____. (1984). *The Significance of Philosophical Skepticism*, Oxford: Clarendon Press.

_____. (1989). "Understanding Human Knowledge in General," in M. Clay and K. Leher (eds.), *Knowledge and Skepticism*, Bolder: Westview Press, 31-50.

Tichy, P. (1986). "Putnam on Brains in a Vat," *Philosophia*, 16: 144-145.

Tymoczko, T. (1989). "In Defense of Putnam's Brains," *Philosophical Studies*, 57: 281-297.

_____. (1990). "Brains Don't Lie: They Don't Even Make Many Mistakes," in M.R. Roth and G. Ross (eds.), *Doubting: Contemporary Perspectives on Skepticism*, Dordrecht: Kluwer, 195-214.

Weitzman, L. (1996). "What Makes a Causal Theory of Content Anti-Skeptical?" *Philosophy and Phenomenological Research*, 61.
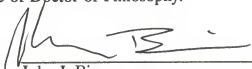
BIOGRAPHICAL STATEMENT

Gayle Brown received her B.A. in psychology and philosophy from George Mason University in 1991. She received her M.A. in philosophy from the University of Florida in 1994. She is expected to receive her Ph.D. in philosophy from the University of Florida in 1999.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.
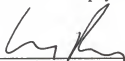
Kirk A. Ludwig, Chairperson
Associate Professor of Philosophy

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

John I. Biro
Professor of Philosophy

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Gregory Brian Ray
Associate Professor of Philosophy

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Chuang Liu
Associate Professor of Philosophy

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Vassiliki B. Smocovitis
Associate Professor of History

This dissertation was submitted to the Graduate Faculty of the Department of Philosophy in the College of Liberal Arts and Sciences and to the Graduate School and was accepted as partial fulfillment of the degree of Doctor of Philosophy.

August 1999   _____
                                     Dean, Graduate School